Involving Practitioners in Evaluation Studies:
How Viable is Collaborative Evaluation in Schools?

Jean A. King

Center for Applied Research and Educational Improvement (CAREI)
College of Education
University of Minnesota

March, 1993

The notion of involving practitioners in a school-based research process is not new. John Dewey and his colleagues at the turn of the century gave teachers and students direct roles in ongoing inquiry, and the progressive tradition, with newfound support from the teacher researcher movement (Stenhouse, 1975), has continued to encourage that practice in schools. Action research (AR), named by social psychologist Kurt Lewin in the late 1940's, brought together university-based researchers with community individuals to engage in collaborative problem-solving around some of the most difficult issues of the day (e.g., racism). The related rise of educational action research in the 1950's (Corey, 1953; Taba & Noel, 1957) pointed to the potential of research collaboration to effect meaningful change in schools, but a number of factors—e.g., the lack of time for such work in the traditional school day, methodological challenges from the research community, and the launch of Sputnik followed by the development of a federally-funded RD&D model—led, until recently, to its decline in this country (King & Lonnquist, 1992). The currently burgeoning literature on teacher research (e.g., Kincheloe, 1991; Cochran-Smith & Lytle, 1993) and educational action research (e.g., Holly & Whitehead, 1984) points to the re-emergence of this process as a means of professional development, school improvement, and, some (e.g., McTaggart, 1991) would claim, long-term social change.

In contrast, the field of program evaluation, developed in part to meet the evaluation needs of the federally funded RD&D model, has traditionally assigned school-based evaluation responsibilities to others. However, it is clear that this distinction is being eroded in many schools, as practitioners are being encouraged to take on roles that were once the sole province of the professional. This shift has been driven by a number of factors, including the recognition that practitioners are often the best equipped to make decisions about the effectiveness of educational programs, and the desire to involve practitioners in the improvement of schools.

The conceptual distinction between the terms research and evaluation is important, and the processes that I am discussing here should properly and technically be called evaluation. However, throughout the paper, I use the two terms interchangeably in part because “action research” includes the word research, and in part because I am taking linguistic liberty for the sake of variety in my language.

Contrary to common belief, action research did not, however, disappear. As Noffke (1990) documents, action research work continued in a number of fields.
educational practitioners a different role. Stereotypically, they are first the data sources and last the eventual recipients of the products of evaluation. Whether or not anyone uses these products—and what it might take to get someone to do so—has been a topic of discussion in the evaluation utilization literature for well over a decade (see, Alkin, Daillak, & White, 1979; King & Pechman, 1982; Patton, 1986). Fairly early on in the field, the notion of responsive evaluation (Stake, 1975) suggested the importance of responding to the information needs of people participating in an evaluation study, an approach extended in the stakeholder model to include virtually all possible participants (Cousins & Earl, 1992). Patton’s utilization-focused approach (1986) created an evaluation process that was systematically structured around the questions of those who would use the results, an ongoing and interactive collaboration between evaluator and practitioner. Participatory evaluation (Cousins & Earl, 1992) makes this collaboration explicit. It is an applied social research that involves a partnership between trained evaluation personnel and practice-based decision makers, organization members with program responsibility, or people with a vital interest in the program (pp. 399-400).

Over time, successful participatory evaluation seeks to transform schools into learning organizations, building their research capacity to go it alone, without an outside evaluation collaborator. “We see participatory evaluation as a powerful learning system designed to foster local applied research and thereby enhance social discourse about relevant organizational issues” (p. 401).

The purpose of this paper is to discuss the related concepts of action research and participatory evaluation, first conceptually and then through brief case descriptions, two of participatory evaluations and one of a collaborative action research effort. Personal lessons learned across these approaches will be presented, followed by a short conclusion.
A Comparison of Collaborative Evaluation/Research Approaches

Figure 1 compares four evaluation/research approaches that involve practitioners. Two--the stakeholder model and participatory evaluation--are explicated in Cousins and Earl (1992). The remaining two--traditional and practitioner-centered action research--are presented in King and Lonnquist (1992). Definitions suggest that these approaches share two defining characteristics: they actively involve practitioners in the research process; and they study existing programs in situ. With the exception of the practitioner-centered action research category, their differences are more of degree than kind:

- Practitioner involvement in the stakeholder approach, although important, may be fairly minimal. In a stakeholder evaluation, evaluators purposefully involve people who have a stake in the program--as administrators, policy makers, clients, service deliverers, etc.--and work throughout the process to insure that, to the extent possible, the needs and interests of multiple audiences receive attention.

- In traditional action research, a small number of practitioners are involved, providing input to researchers (typically university-based), making suggestions, facilitating data collection, and so on. Practitioners play an active role, but outside researchers insure the validity of a research process that is designed to generate social science data while simultaneously addressing real problems.3

- The role of practitioners in participatory evaluation is also active. In fact, a small number of people who will become the primary users of evaluation information engage in the very “nuts and bolts” of the process. In contrast

3 The concept of participatory action research as practiced in the field of organization development in business, industry, and agriculture (see, e.g., Whyte et al., 1989; Whyte, 1991) uses the process of traditional action research.
**Figure 1: A Comparison of Collaborative Approaches to Evaluation**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Stakeholder</th>
<th>Traditional Action Research</th>
<th>Participatory Evaluation</th>
<th>Practitioner-Centered Action Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who’s in charge of the study?</td>
<td>Evaluator as principal investigator</td>
<td>Researcher/evaluator, with input from practitioner</td>
<td>Evaluator, with assistance of practitioner/student</td>
<td>Practitioners, with assistance of research consultants</td>
</tr>
<tr>
<td>What practitioners are involved in the study?</td>
<td>Large number of stakeholders; any-one with a stake in the program</td>
<td>Small number of those who actively engage in the study; people who are interested or helpful</td>
<td>Small number of primary users</td>
<td>Action researchers</td>
</tr>
<tr>
<td>How do practitioners participate?</td>
<td>Consultative role; give information on context, information needs</td>
<td>Active role; ask questions, give input, help analyze, etc.</td>
<td>Active role; engage in &quot;nuts and bolts&quot; of evaluation process</td>
<td>Active role; control the research process</td>
</tr>
<tr>
<td>What time-frame?</td>
<td>Length of study</td>
<td>Length of study</td>
<td>Length of study OR Ongoing</td>
<td>Ongoing research cycle; organizational learning</td>
</tr>
<tr>
<td>What involvement in theory development?</td>
<td>[If any] Develop program theory</td>
<td>Create traditional social science theory</td>
<td>Develop program theory (theories-in-action, mental models)</td>
<td>[If any] Develop practical theory OR critical theory</td>
</tr>
</tbody>
</table>
to traditional action research, participatory evaluation has long-term implications for the practitioners it involves. Over time, the evaluator would hope to teach site collaborators sufficient research skills to develop organizational learning within the organization. In that sense, the goal of a participatory evaluator is eventually to put him- or herself out of work when the research capacity of the organization is self-sustaining.

- The important difference in practitioner-centered action research is that the researcher/evaluator no longer is in charge of the evaluation process. Instead, site-based practitioners themselves engage in a five-step process through which they frame their own questions, develop their own designs, collect and analyze data, reflect—and then begin the collaborative cycle again. Outside facilitators can join in the process, but only to provide expertise as “critical friends” and peers, not as controlling superiors. In practitioner-centered AR, it is more important to answer the right question however messy the methodology, than to answer the wrong question extremely well, and issues of utility and feasibility sometimes demand trade-offs in technical accuracy.

These four approaches each give program practitioners a role to play in an evaluative process. I would argue that the concept of organizational learning upon which Cousins and Earl (1992) base their theoretical justification for participatory evaluation at best is exactly what takes place in practitioner-centered action research. But what exactly is the role of evaluator in these collaborative approaches?

Like Mollière’s bourgeois gentilhomme, who proudly realized that he had been

---

4 I am speaking here primarily of what critical and emancipatory action research theorists (e.g., McTaggart, 1991; Kincheloe, 1991) would call practical action research. Emancipatory action research, developed from the perspective of critical theory, uses a broader frame for its work than I am applying in this discussion.
speaking prose all his life, I was rather pleased recently to realize that I have engaged in participatory evaluation virtually my entire career. Having identified my modus operandi, I became interested in reflecting on my practice across three studies, each of which created practical challenges. Before describing the lessons, I must briefly describe each case.

Three Program Examples: CPIP; OEL; and CARP

Extending over the past 14 years, my involvement in collaborative evaluation approaches has included three projects that provided heartfelt lessons. The first two are examples of what I would now call participatory evaluation; the third is a current project involving action research.

CPIP. The Curriculum Process Improvement Program (CPIP) was an effort to place a learning for mastery (LFM) approach (Bloom et al., 1971) into the grade 7-12 curriculum of a large city school district. CPIP was an expensive program. Benjamin Bloom himself gave the keynote during a highly public program launching in the late 1970's. Unit by unit, CPIP tests were created in various subjects. Centralized computer scoring was supposed to provide rapid feedback to teachers and students, as well as create a district record for accountability purposes. Structurally, a group of teachers was released from classroom responsibilities to serve as district CPIP trainers, and all secondary teachers were eventually cycled through a two week training. District administrators proudly prepared transparencies and audiotapes proclaiming the success of CPIP.

But as is often the case when victory is proclaimed before all battles are fought, the CPIP implementation ran into major snags. While the trainers were extremely enthusiastic about the LFM process, those they trained were often less committed. Some questioned the viability of a mastery learning approach in their content area. Others resented the amount of 

---

5 Certain details related to the cases have been altered to conceal the identity of specific individuals or sites. The basic content has not, however, been changed.
work required to keep the program going with over 150 students a day, even assuming the computer support worked smoothly, which, from the beginning, it did not. Teachers who dutifully submitted their CPIP computer sheets often waited weeks for the scores to return, and, by then, the effectiveness of the feedback was at best limited. Stacks of user-unfriendly test results sat on teachers’ desks across the district, and word quickly got out that this was a program in trouble. Despite major expenditures for staff development and computer support, CPIP wasn’t working.

Within three years, exit the superintendent who created CPIP, and enter another, who immediately hired a go-getting deputy superintendent for curriculum to shape things up in that area. In an era of tightening budgets, CPIP’s price tag made it an extremely visible target for cuts. The well-known problems related to computerized scoring and classroom implementation provided no evidence to support continuation, and rumors circulated that the deputy had set his sites on CPIP. However, because he was new and did not want to appear unfair, the deputy superintendent commissioned the district evaluation office to conduct a full-scale CPIP evaluation study. The evaluation director, well aware of CPIP implementation realities, was committed to an inclusive process that would create useful information about the program and its outcomes, regardless of its eventual fate. Knowing the professional commitment of the staff developers, she proposed a participatory effort through which their concerns would be joined with those of the deputy superintendent. He refused to collaborate, telling the director that the evaluation was her job and questioning why his involvement was even an possibility.

Suffice it to say that the evaluation was ill-fated from the beginning. People across the district saw the writing on the wall, so the incentives for participation were few. Nevertheless, some CPIP trainers did collaborate over the course of the year, attending meetings, brainstorming issues, and agreeing to questions and methods. Their hope was that their hard work and good efforts could somehow be salvaged and that the program
would continue in some form. Meanwhile, classroom teachers dutifully completed questionnaires, documenting their frustration with the program, which still was not working well.

The CPIP bombshell was dropped shortly after the evaluation director informally presented the initial results of the study to the deputy superintendent. During the budget planning for the following year, CPIP funding was noticeably absent, and, when questioned, the deputy cited the evaluation outcomes as the rationale for the cut. The CPIP staff developers were devastated. Their involvement in the evaluation had done nothing to save their central office positions, and most were reassigned to secondary classrooms. Remnants of CPIP lived on in some of these, but no one in the central administration used the data that detailed what had been learned from this massive experiment in LFM to make decisions about a new secondary curriculum.6

OEL. The Office of Educational Leadership (OEL) existed in the Minnesota State Department of Education for two years (1989-1991), funded by the state legislature to build the foundation for the radical transformation of schooling in Minnesota. Enabling legislation directed the OEL to develop a two-year research and development project to determine the effectiveness of an outcome-based system of education, and develop it they did. The four person OEL staff set about to do nothing less than re-design the process of education at ten school district project sites, each funded with a grant of $50,000 a year for two years. In September, 1989, the ten sites, comprised of 17 districts and five educational consortiums and representing rural, suburban, and inner city communities, began what they were told would be a long-term change effort (10 year plus) that the legislature would fund.

In March, 1990, the OEL contracted with our organization, the Center for Applied

---

6 As a poignant commentary on the situation, the deputy superintendent left the district in a year and a half, saying that the district politics made his job unworkable. The CPIP trainers reflected that they might have still had their program and jobs had he not been hired.
Research and Educational Improvement (CAREI), to lead a collaborative research process that would model the kind of organizational learning that schools of the future might use. This process was based on the assumption that collaborative problem-solving involving—at some level—all OEL participants would facilitate meaningful transformation at the project sites. This was participatory evaluation writ large, with the expectation that over time, sites truly would take on their own research. Initial research planning included over 50 individuals, representing each of the ten sites, CAREI, the Minnesota Department of Education, and the College of Education at the University of Minnesota; we did not, however, invite legislators or their aides. Michael Quinn Patton facilitated a two-day planning session in April, 1990, which resulted in guiding principles for our collaborative research, baseline data questions, and a list of appropriate methods. Much to my pleasure and amazement, we left with virtual consensus both on a participatory evaluation process and on the questions the group wanted answered.

If that was the momentary good news, the longer-term bad news became immediately apparent. It was mid-April, and year 1 baseline data, including classroom observation data, had to be collected before schools let out in June. Knowing how schools wind down, this actually gave us about four weeks to gain entree at ten sites across the state, which were actually 37 or 39 different schools, depending on how you counted them. In retrospect, the frenzy of that data collection was somewhat reminiscent of piranhas feeding on a hapless beast. Fifteen graduate students descended on project sites with a fairly long list of data requirements. While the list had been developed collaboratively in the

---

7 The guiding principles, revised in September, 1990, were as follows: the research process must respect the uniqueness of the learning sites; it must be supportive of learning without being a burden to learning sites; it must involve, receive input from, and be sensitive to multiple levels of diverse stakeholders; it must include multiple and “people-sensitive” methods; it must foster communication and understanding of the ongoing change process; research outcomes must be useful within and across learning sites; and they must provide tangible feedback on student success and solid, usable data about what is good for learners (King et al., 1991).
planning session, people at the sites were often unaware of what their colleagues had agreed to, and, as the end of the year crunch set in, they were sometimes frustrated by the demands of this state-funded, university-personed evaluation process. Some CAREI staff added to the problems. One research assistant inadvertently offended a key figure at one site, who basically cut off data collection for the year. Another RA, in the throes of a serious illness (we later learned), exhibited strange behavior that made people who interacted with her question the evaluation process. Nevertheless, the team managed somehow to collect all the needed information and then set about writing the 39 short case studies summarized in Volume II of the OEL Phase I Evaluation Report, a report in which we took great pride.

Determined never again to put ourselves in such a position, we convinced the OEL Director to create a group of site representatives for ongoing discussion and project monitoring. The OEL network, with representatives from OEL, CAREI, and each site, began meeting in September, 1990. An early decision of the group was that a short survey of people involved in the OEL effort would provide useful information for the legislative session beginning in January. That was done and, with the cooperation of people at the sites, completed with less stress than the preceding year. The network then planned more extensive work for the spring, including a core study across sites and the option of a voluntary site-specific study. With fewer and more reliable RA’s, a more reasonable timeline, and a lot of help from our on-site collaborators, the Phase II data collection--individual interviews with administrators and/or opinion leaders; group interviews with site leadership, teachers, parents, and students; and a survey of all administrators and instructional staff across all sites--was completed in a timely manner. We were truly confident that the participatory evaluation process was succeeding. Certain principals, central office administrators, and even some classroom teachers were helping frame questions, administer surveys, and make sense of the information with us. These individuals were becoming our colleagues and friends, and the process of watching the
OEL effort grow was invigorating.

We should not have been so sanguine. While we did our work, the recession hit Minnesota, dropping state revenues drastically. The legislative protectors of outcome-based education and the OEL left the legislature, leaving the OEL Director in a vulnerable position. Changes in the Department of Education made the OEL a virtual step-child, and rumors began circulating across sites that funding trouble was brewing. The ten year development process we had taken as a given suddenly vanished, and there was talk, to our dismay, that after two years of intense effort, the OEL would not be reauthorized. Mazzoni et al. (1992) document in detail the fate of the OEL, including the one and only clear impact of the fruits of our evaluation labor. The sole remaining legislator who had been head of an education committee when the OEL was first created commented to a group of OEL site representatives that the Phase I report was a “piece of ---” because it included no standardized test scores. This was one stated reason for recommending the program’s elimination.

It is important to note that an explicit outcome of the April, 1990, planning session was the group’s consensus that it would be wrong to include test scores in the early evaluation of the OEL transformation process. To do so would send an inappropriate signal to school sites that our study represented evaluation business as usual. If you’re serious about transformation, the argument went, then you also need to transform the evaluation process. In retrospect, we all questioned our seeming naivete and the wisdom of our collective decision. With the dissolution of the OEL office, the evaluation process ended.

CARP. The Collaborative Action Research Project (CARP) is an umbrella that encompasses a number of practitioner research activities fostered by CAREI. For the purposes of this paper, I will discuss the practitioner-centered action research we have been
studying for two years. Three teams of teachers--two from elementary schools, one from a high school; two in outer-ring suburbs, one in the inner city--have collaborated around an important change issue in their building. Their issues are the following:

- The urban high school site is applying the principles of the Coalition of Essential Schools to develop and assess higher levels of trust and respect within their team’s students and teachers. A number of actions were taken, e.g., teachers taught specific listening skills and peer mediation techniques, they modelled appropriate behavior, and students helped to design a team governance structure. Pre-, post-survey data, focus group interviews, and anecdotal evidence suggested some success by June of the first year. As a result, the team committed to continuing study of this topic in the second year.

- One elementary site is studying multi-age grouping (MAG). In year one, they conducted a MAG experiment. The team developed ten one-hour lessons around the theme of the egg to integrate children from two or three grade levels in innovative, MAG activities. To their surprise, 12 other teachers agreed to participate by sharing children and teaching the lessons to mixed age groups. While the results of the experiment were also mixed, the action researchers became more fully committed to the notion of multi-age grouping and to working as a close-knit team. Fortunately, their new principal supported their proposal this year to have two first grade classes and a second grade class, so they are studying their continuing efforts to

---

8 This work is funded, in part, by the Center on Organization and Restructuring of Schools, supported by the U.S. Department of Education, Office of Educational Research and Improvement (Grant No. R117Q00005-91) and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison. The opinions expressed are those of the authors and do not necessarily reflect the views of the supporting agencies.
create interactions across grades and to use developmentally appropriate curriculum.

- At the third site, an early childhood center, the team began last year by studying the effects of mainstreaming on special needs children in their classes. During the course of their work, the action researchers innocently ran up against district special education politics related to pull-out vs. inclusionary programming and learned lessons that went well beyond their classrooms. Nevertheless, they concluded year one with information that would help individual children as they left the center as well as continuing practice with new children, and they believed that the district central office understood their concerns. Their work this year has switched focus to another special population—children needing ESL instruction.

Our role as evaluators/researchers in this practitioner-centered process has been an ongoing learning experience. Unlike the other two cases where I was clearly in charge of the evaluation process, as an action research collaborator my job is to facilitate the teachers’ ongoing research. Knowing how busy they are and understanding that they have lives outside of work, this can be frustrating and discouraging. Our staff sets up meetings, buys the snacks (not a trivial detail), takes notes on who’s doing what, and, in the words of more than one researcher, serves as “cheerleaders” for the process. We do render technical assistance, but unobtrusively, and we do try to intervene positively when problems arise. For example, when teachers at one site were concerned about being allowed to participate in an important meeting during school hours, we cleared their involvement through the central office. One thing is clear: after our first year together, the teams this year are far more independent, and our role as outside consultant has diminished considerably. This “success” feels strange as teams tell us they no longer need us. We are turning now to other teams and working to spread the action research process farther into the practice of
Minnesota’s schools.

Lessons Learned from Collaborative Evaluation/Research Studies To Date

Sporting tatters and bruises is not necessarily a sign that someone has learned something. Over the course of time, some people have the same experience many times, while others analyze their experiences to grow and develop. I hope I fall in the latter category. The three studies described previously created dynamic classrooms\(^9\) for my evaluation practice, and reflection on this work points to four lessons: the importance of people’s wanting to participate; the power of political context; the necessity of continuing support; and the need for common meaning between distinct worlds of practice.

Volunteers needed. To paraphrase an old saw, you can lead a person to evaluation, but you cannot make him or her participate. The power of the human factor has long since been accepted in the evaluation literature (e.g., Patton et al., 1978), but the special power of people to affect collaborative studies both positively and negatively cannot be overstated. My evaluation experience documents three types of individuals in schools. Fortunately, there is one group of individuals eager to join forces with evaluators in collaborative efforts. These individuals become primary users and potential creators of evaluation information.

A second group are those who, for various reasons, are unable to engage in meaningful collaboration. Some reasons are obvious. For example, in the action research study, two teachers were unable to collaborate for health reasons; halfway through the process, one developed a partial paralysis, and the other fell and broke his foot. Another teacher had small children and little time for projects that extended the work day. In the second year of the study, a new principal wanted to be involved, but, given the challenges of learning a new faculty and program, could only participate on an occasional basis. Other

\(^9\) Actually, the image of fiery crucibles immediately sprang to mind!
reasons making people unable to collaborate are less obvious. Control issues seem important. In the OEL study, for example, one of the state department collaborators early on became concerned that the research agenda was slipping from her fingers and, verbal support to the contrary, that the translation of the language of collaboration into action had simply gone too far. A tense confrontation resolved the issue at least temporarily, but the evaluation team, blindsided once, never again fully trusted that individual.

A third category of people are those who choose not to participate. In the CPIP study, for example, the deputy superintendent who commissioned the evaluation wanted only the results--and he had a clear sense of what his ultimate decision would be even before one piece of data was collected. He was neither interested in collaborating nor willing to interact with those conducting or affected by the study. Given his status in the district, he could opt to do what he wanted. The OEL project also included people who chose not to engage. Six of the ten sites elected not to develop a site-specific evaluation, and some did not actively participate in the network meetings that determined the course of the study. Clandinin and Connelly (1992, p. 377) note that requiring teachers to become emancipated is itself a form of domination, and collaborative approaches simply cannot empower people who are not willing to be empowered.

Collaboration is not just a set of skills to be taught. The process demands a commitment that some people are unable or unwilling to make, and evaluators must accept these individuals’ choices. Practitioners must also acknowledge that not all evaluators will choose to engage in participatory practice. Perhaps over time, if collaborative approaches become more common, best practice will necessarily include some form of meaningful interaction between evaluator and primary user. In the meantime, requiring people--whether practitioners or evaluators--to participate in an evaluation process seems an unlikely way to insure its success.

The political context. My experience has taught me that the collaborative evaluation
process is fragile, with virtually continuous opportunities for failure. Prime among these is the politics of evaluation settings, which can destroy a collaboration in a single thoughtless comment. To fail to acknowledge this fact is to risk setting people up for extreme frustration. Each of the studies described provides an easy, if sad, example of the truth of this claim.

- In the CPIP evaluation, the training coordinators willingly worked with the R&D evaluator, hoping that positive documentation would both justify the program’s value and enable at least some of them to keep central office staff development positions. When the deputy superintendent eliminated the program with seemingly little attention paid to the evaluation data, the coordinators and the evaluation staff felt betrayed. The district evaluation head wondered, in retrospect, if the collaboration had been worth the upset it unavoidably created.

- One of the most frustrating moments of the OEL study was when we learned that the legislator mentioned earlier cited our study, which purposefully contained no standardized test scores at the end of the first year, as one of the reasons for cutting the program. Mazzoni et al. (1992) suggest that this individual had his own reasons for wanting the program eliminated, but for school district collaborators, the attack on the research suggested an egregious error on our part. Their efforts to provide existing test scores to the legislator in question were rejected in a face to face meeting that left many insulted or furious.

- Building-level politics have played an interesting role in our action research efforts. The teacher researchers report feeling like “elite lepers,” rebuffed, on the one hand, because others see them as the chosen few who garner special favors from the administration, yet ostracized, on the other hand, because
they appear overly enthusiastic about school improvement and professional
development. At one site, for example, to continue the action research
process is to risk the disfavor of an experienced faculty who view the
researchers as overly committed and idealistic do-gooders. As outsider
facilitators, we cannot protect our research collaborators, but we can at least
engage them in reflective dialogue and a professional community. At the end
of last year, one teacher said, “We were validated as professionals for our
own observations”; this made us feel validated as well.

The point to be emphasized here is that sometimes--and not surprisingly--the politics
of a situation can overwhelm even the best of collaborative intentions. Whether or not to
proceed with a participatory process is not a decision to be made lightly. My experience
suggests that it is better not to begin if the politics will unavoidably limit the likelihood of a
successful collaboration. For example, the CPIP task given the evaluation unit was clearly a
hatchet job; asking people to participate in the demise of their own program did not, in the
end, make them feel better about its elimination. A second example comes from the OEL
study. Through no one’s fault, but owing to political demands for accountability, its first
year carried an impossible timeline; the planning could not begin until early April, which
meant that we collected data in May and even as late as June--predictably the worst months to
ask teachers, students, or administrators to reflect quietly on their experiences. School-based
people blamed university collaborators for their lack of sensitivity to the facts of school life,
and in retrospect we might have built a stronger project had we convinced our state
department collaborators to wait until the fall.10

The first two lessons suggest necessary conditions for collaborative evaluation:
finding people who want to participate in a political setting that makes participation viable.

10 Ultimately, of course, this would have made no difference in the program’s demise.
Assume that you’ve led willing practitioners to the evaluation project, ready to jump in. Assume also that the politics surrounding the project even make the process possible. The third and fourth lessons speak to what the evaluator must do to grow the ongoing research connection.

**Building support and trust.** My experience suggests that even willing collaborators will limit or drop their involvement if they do not feel supported during the study or they do not trust the people involved. Evaluators who collaborate with practitioners need to pay considerable attention to motivating and supporting their involvement from beginning to end, and trust building must be ongoing. In part this can be done structurally. By making regular connections a structural part of the evaluation process and two way communication an ongoing reality, evaluators can create necessary venues for discussion and reflection, both on the content of the study and on the collaborative process itself. Informal connections should also be systematically included so that different individuals have a number of possibilities for expressing their concerns.

Structure is important, but support and trust building is also a matter of belief. Evaluators must project an attitude of true interest in practitioners’ concerns, respect for the challenging world of school practice, and a continuing openness. No problem is too big or small to be discussed, and, to the extent possible, it is the evaluator’s job to resolve glitches when they arise. One of the bigger challenges for me personally comes from the fact that, as a collaborative evaluator, I never get to be whiny and small-minded, even when people are systematically undermining a carefully wrought study. With each person added—whether evaluator or practitioner—the trust building process begins anew. As the mother of the process, I must be the calm person above such pettiness, working tirelessly to make it succeed. Wearing a constant smile can be tiring. Reaching a higher level of trust, where I can truly express my feelings and opinions about the evaluation process, has been relatively rare. Consider the following examples:
Providing support to site-based collaborators was a problem throughout the OEL study. In the first year, there was an impossible timeline and simply too many of them and not enough of us to do more than field complaints. In the second year the planning network created a smaller group, but it met only occasionally and not everyone attended every meeting. Connection with state department colleagues was better, but because the staff had literally taken on the task of transforming education at the project sites (along with teacher education in the state), they had virtually no time to discuss evaluation issues until there was an external pressure or minor crisis that required resolution. Although we had good working relations with virtually all the sites, the extreme lack of trust evidenced by the confrontation at the state department described earlier suggests the degree to which our overall collaboration failed.

CARP provides several counterexamples to this, cases where our ongoing research connection has resulted in a growing and viable partnership (as well, I would add, as personal friendships). At one site, for example, a team of teachers refers to us as “friendly nags” because when we call to schedule a meeting, they feel that they have to have something new to report and this motivates them to take the next step in their study. To the initial surprise of another group, we have met teachers at their school at 7:15 in the morning, in the afternoon, or whenever they are available, trying to fit the action research into the limited time slots they have available. The CARP group as a whole chuckles at our “meeting and eating” together, but over the course of two years, the feeling of support and trust generated by what Huberman (1990) would call this reciprocal engagement has been motivating to us all. When we needed questionnaires completed at each school site on fairly short notice, our collaborators didn’t blink an eye; they willingly took the surveys,
administered them, and got them back to us.

Creating a common meaning. Even the existence of support and trust is not sufficient to insure an effective collaborative process. In addition, collaborative evaluation efforts demand that practitioners and evaluators truly speak the same language in order to make collective sense of the evaluation process and its outcomes. For some, this requires changing the common belief (and, in many places, the virtual reality) that practitioners and evaluators live in distinct worlds. For teachers and school administrators, the evaluation process may seem irrelevant, time-consuming, and overly quantified. As Jones (1989, p.51) notes, its results may have the "cutting edge of sponge" in the eyes of practitioners. For their part, evaluators may be concerned with the immediacy of daily practice and the messiness of trying to study it. I continue to be surprised when school people think I'm selling collaborative evaluation and action research for my own benefit, i.e., that somehow it's of value to me as a educator, but not to them. In part this stems from a tradition that places researchers and evaluators above school-based personnel on some grand hierarchy of education, although the growth of teacher research (see, e.g., Cochran-Smith & Lytle, 1990; Kincheloe, 1991) suggests that this may be changing.

In my evaluation experience, people have placed me in three distinct roles. Some want me to be their hired hand, disappearing after being given an assignment and reappearing at some point with a useful analysis--or at least a report to file somewhere. A related role is that of the outside researcher, the dispassionate "scientist" who lends credibility to the evaluation process and its results, often in contentious situations. "She's from the university," people whisper even when I leave my lab coat at home. The third role I am asked to play--and the one I actively cultivate in my practice--is that of critical friend. This notion, taken from the action research literature, suggests an objective participant in an ongoing discussion of discourse, activities, organizational structures, and values: in short, a true collaborator in long-term program improvement. The roles of internal evaluators and the
notions of organization development, developmental evaluation (Hallie Preskill, personal communication, January, 1993), and perhaps even the emancipatory evaluation that will be the theme of the 1993 meeting of the American Evaluation Association suggest that this is not usual. For people seeking someone to fill the first two roles, however, acceptance of the third may require some active explanation and instruction over time—if, and only if, they are willing to learn.

How is the notion of shared meaning created in practice? The studies provide examples of how we have worked to create the common ground of mutually acceptable, shared meanings with our site-based colleagues. We work to develop and maintain primary users.

• Throughout the two year collaboration, the OEL study sought to create a shared world by documenting the situation in which the schools found themselves—forced to implement top-down, intrusive mandates; inhabited by challenging students different from those of earlier years; assailed by parents who assume the schools are either “not broken” or in need of major reform; and running hard just to stay barely in place (King & Bosma, 1991; King et al., 1992). Some school-based participants reported pleasure when the first year report passed their litmus test by including negative information about the project. State department staff were predictably less pleased!

• In addition, our team struggled with site-specific tensions (e.g., issues related to two communities’ long-term racism and hostility towards “outsiders”) and sought to infuse evaluation information back from whence it came. The shared meaning in one case was welcomed; district staff included rather negative OEL community data in a grant proposal for funds related to improving diversity. In the other case, evaluation participants agreed that while racism was, indeed, a problem affecting the school, it was an open
secret that could not be made part of the evaluation record.

- Our involvement in action research provides the most visible examples of a merging of worlds as the cycle of planning, acting, observing, and reflecting places teachers directly in the practice of evaluation. The teacher researchers have become the immediate beneficiaries and users of the information they generate, and they often report their surprise at the relative straightforwardness of the process. In the words of one of the teachers, “Action research helped us to focus on the problem and bring awareness to the entire building staff. It provided a framework or a process to work through--to come to some understanding of and some solutions to these problems.”

As is often the case with lessons learned through experience, the four that I discuss here are not surprising. I probably could have created the list after reading the tenets of good evaluation practice. But that is also the reason they are important. If we are serious about nurturing practitioner involvement in evaluation studies, I would argue that we must pay direct attention to potential participants, likely political contexts, the development of ongoing support and trust, and the creation of shared meaning.

Summary and Conclusion

Are collaborative evaluation approaches viable in school settings? I would answer an enthusiastic maybe. In this paper I began by comparing four conceptual approaches, each of which involves school-based practitioners. I then described three cases that used a participatory process, followed by the lessons I learned from those experiences. In two of the cases (CPIP and OEL), the evaluation process was unable to counteract the powerful political setting within which the programs operated. This is not entirely a fault of the process, but at one level the outcomes were nevertheless damaging to those who agreed to participate. While we all learned from each other and the events, in both cases people would
have been reluctant to commit quickly to a similar research involvement.

To date, few existing school organizations engage in organizational learning of the type Cousins and Earl (1992) discuss. We need research evidence that this process is worth the effort, and, based on my ongoing study and reflection as an evaluation practitioner, I have hope that the developing practice of practitioner-centered action research may provide us such evidence. The four lessons I cited are all applied in an action research effort: only volunteers take part in the process since they must study themselves and take charge of research activities; action research takes place in a known context, and, while you may still have political challenges, the probability of being blindsided is greatly decreased; support and trust are characteristics of the professional community fostered by action research; and the process itself brings research to life in the school setting, thus merging the worlds of research and practice. Numerous questions remain, relating, e.g., to the training necessary for such efforts, the role of theory development in action research, and the effect of such research on teacher behavior and student learning. From my perspective, however, the prospects for continued development of collaborative evaluation approaches at this point look bright.
References


Cochran-Smith, M., & Lytle, S. (1993). Inside/outside:


