Developmental Evaluation

MICHAEL QUINN PATTON

INTRODUCTION

Develop. Synonyms: Improve. Evolve. Create. Form. Grow. Mature. Unfold. Ripen. Change.

Language matters. It simultaneously suggests possibilities and communicates boundaries. We can't think without it and with it our thinking is inevitably constrained. Why this reminder from Communications 101? Because what we call what we do affects what we do and how we do it—defines it to ourselves and others.

Edward Sapir studied "The relation of habitual thought and behavior to language." He concluded:

> Human beings...are very much at the mercy of the particular language which has become the medium of expression for their society...The fact of the matter is that the "real world" is to a large extent, unconsciously built up on the language habits of the group...We see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation. (Rheingold, 1988, p. 11.)

Evaluators constitute a subculture. Our language habits predispose certain choices of action and interpretation. As we attend increasingly to issues of diversity, multiculturalism and global perspectives, our language habits and corresponding predispositions challenge all that we do.

EMPOWERMENT EVALUATION

The provocative phrase "empowerment evaluation," to many an oxymoronic rendering, opened up new possibilities as the theme of the 1993 American Evaluation Association national conference. It's premature to predict if the phrase will endure, but I talked with several first-time attendees who said they'd never thought of evaluation as potentially empowering. They expressed curiosity, confusion, wonder, hope, and skepticism. Some

Michael Quinn Patton, Utilization-Focused Information and Training, 3228 46th Ave. South, Minneapolis, MN 55406-2329.

old-timers dismissed the phrase as faddish, meaningless, politically corrupt doublespeak, and even—Can you imagine?—values-laden. Evaluation? Values-laden? I was naturally shocked at the suggestion. And I realized, right there in Dallas, Texas, that the world will not end in a subjective bang, but in a boring whimper as voices of objectivity drifting off into the chaos. The ultimate development.

Which brings me back to my theme: developmental evaluation. I have begun using this phrase to describe certain long-term, partnering relationships with clients who are, themselves, engaged in ongoing program development. I have found these relationships to be substantially different from the kinds of evaluations I conducted earlier in my practice. And more personally satisfying. My role has become more—How can I describe it?—*developmental*. I'll explain what that means, but first…

A BIT OF HISTORY AND CONTEXT

Evaluation, like the urban poor, grew up in the projects—large government structures aimed at concentrating limited resources on seemingly unlimited problems. House (1993) has insightfully reviewed the emergence of evaluation as a profession under the tutelage and hegemony of government. Regardless of political-economic system, the dominant motif of modern government worldwide is planning. That means assessing needs, identifying solutions, targeting populations, setting goals, procuring resources, implementing programs, and, of course, evaluating results. It's all very logical and rational.

Of course, imperfections exist: inadequately assessed needs; fuzzy and conflicting goals; poorly defined targets; insufficient resources; and sloppy implementation. But those are precisely the problems evaluations expose. The logic remains intact.

Within this logic, we've distinguished two kinds of evaluations: summative and formative. Summative evaluations judge merit and worth: the extent to which desired goals have been attained; whether measured outcomes can be attributed to observed interventions; and the conditions under which goals were attained that would affect generalizability and therefore intervention dissemination. Formative evaluations help programs get ready for summative evaluation by improving program processes and providing feedback about strengths and weaknesses that appear to affect goal attainment.

Scriven's original distinction between formative and summative in curriculum evaluation made it clear that the purpose of formative evaluation was to get ready for summative evaluation. Over time, the meaning of formative evaluation has been enlarged to include any evaluation whose primary purpose is program improvement, where improvement means a higher degree of goal attainment. But connotatively, it seems to me, formative evaluation remains in the shadow of summative evaluation and suggests an inferior, preliminary stage of assessment while the program and evaluation are getting ready for the real thing: summative judgment.

ENTER DEVELOPMENT PROGRAMMING
AND DEVELOPMENTAL EVALUATION

Developmental programs have as their purpose the vague, general notion of development. The process is the outcome. They eschew clear, specific, and measurable goals up-front

because clarity, specificity, and measurability are limiting. They've identified an issue or problem and want to explore some potential solutions or interventions, but they realize that where they end up will be different for different participants—and that participants themselves should play a major role in goal-setting. The process, then, involves engaging participants in setting and achieving their own goals. Program designers observe where they end up and make judgments about the implications of what has happened for future programming and re-engineering. They never expect to arrive at a steady state of programming because they're constantly tinkering as participants, conditions, learnings, and context change. They don't aspire to arrive at a model subject to summative evaluation and generalization. Rather, they aspire to continuous progress, ongoing adaptation and rapid responsiveness. No sooner do they articulate and clarify some aspect of the process than that very awareness becomes an intervention and acts to change what they do. They don't value traditional characteristics of summative excellence such as standardization of inputs, consistency of treatment, uniformity of outcomes and clarity of causal linkages. They assume a world of multiple causes, diversity of outcomes, inconsistency of interventions, interactive effects at every level—and they find such a world exciting and desirable. They never expect to conduct a summative evaluation because they don't expect the program—or world—to hold still long enough for summative review. They expect to be forever developing and changing—and they want an evaluation approach that supports development and change.

Moreover, they don't conceive of development and change as necessarily improvements. Formative evaluation carries a bias about making something better rather than just making it different. From a developmental perspective, you do something different because something has changed—your understanding, the characteristics of participants, technology, or the world. Those changes are dictated by your current perceptions, but the commitment to change doesn't carry a judgment that what was done before was inadequate or less effective. Change is not necessarily progress. Change is adaptation. As one program director put it:

> We did the best we knew how with what we knew and the resources we had. Now we're at a different place in our development—doing and thinking different things. That's development. That's change. But it's not necessarily improvement.

The developmental perspective, as I experience it, feels quite different from the traditional logic of programming in which goals are predetermined and plans are carefully made for achieving those goals. Developmental programming calls for developmental evaluation in which the evaluator becomes part of a design team helping to monitor what's happening, both processes and outcomes, in an evolving, rapidly changing environment of constant feedback and change. These relationships can go on for years and, in many cases, never involve formal, written reports. (I confess that's become a major attraction for me.)

TEAMING AND COLLABORATION

Developmental evaluation isn't a model. It's a relationship founded on a shared purpose: development. What I bring to the design team is evaluation logic, knowledge about effective programming based on evaluation wisdom, and some methods expertise to help set up

monitoring and feedback systems. I become part of the team. I participate in decision-making about the program and facilitate discussion about how to evaluate whatever happens. All team members render evaluation judgments together and decide how to apply the implications of results for the next stage of development.

Four Examples

1. *A Community Leadership Program.* With two other evaluation colleagues I became part of the design team for a community leadership program in rural Minnesota. The design team included a sociologist, a couple of psychologists, a communications specialist, some adult educators, a funder, and program staff. All design team members had a range of expertise and experiences. What we shared was an interest in leadership and community development.

The relationship lasted over six years and involved different evaluation approaches each year. During that time we engaged in participant observation, several different surveys, field observations, telephone interviews, case studies of individuals and communities, cost analyses, theory of action conceptualizations, futuring exercises, and training participants to do their own evaluations. Each year the program changed in significant ways and new evaluation questions emerged. Program goals and strategies evolved. The evaluation evolved. No final report was ever written. The program continues to evolve—and continues to rely on developmental evaluation.

(The original evaluation design called for three years of formative evaluation followed by two years of summative evaluation. By the end of the first year, the idea of summative evaluation had been discarded because the program found the required conditions to be unacceptably constraining, especially holding the program constant during the summative test, reaching consensus on a short list of desired outcomes, and implementing a comparative design. We moved to the developmental emphasis, as partners rather than external evaluators, at that time.)

2. *Supporting Diversity in Schools.* A group of foundations committed to working with the Saint Paul Public Schools for 10 or more years to support multicultural education funded this program. The problem identified by the community was lower levels of success for children of color on virtually any indicator considered. The "solution" called for a high degree of community engagement, especially by people of color, in partnering with schools. The nature of the partnering and interim outcomes were to emerge from the process. Indeed, it would have been "disempowering" to local communities to predetermine the desired strategies and outcomes prior to their involvement. Moreover, different communities of color—African Americans, Native Americans, Hispanics and Southeast Asians—could be expected to have varying needs, set differing goals and work with the schools in different ways. All of these things had to be *developed*.

The evaluation documented developments, provided feedback at various levels from local communities to the overall district, and facilitated the process of community people and school people coming together to develop evaluative criteria and outcome claims. Both the program design and evaluation changed at least annually, sometimes more often. In the design process, lines between participation, programming, and evaluation were ignored as everyone worked together to develop the program. The evaluation reports, written for public accountability, took the form of multiple voices presenting multiple perspectives. These voices and perspectives were facilitated and organized by the evaluation team, but

the evaluators' voices were simply a few among many. (For more on this process and the multivocal approach, see Stockdill, Duhon-Sells, Olson, & Patton, 1992). The developmental evaluation and process are still ongoing as this is being written. No summative evaluation is planned or deemed appropriate though a great deal of effort is going into publicly communicating the developmental processes and outcomes.

3. *A Children's and Families' Community Initiative.* A local foundation made a 20-year commitment to work with two inner city neighborhoods to support a healthier environment for children and families. The communities are poor and populated by people of diverse ethnic and racial backgrounds. The heart of the commitment was to provide funds for people in the community to set their own goals and fund projects they deemed worthwhile. A community-based steering committee became, in effect, a decision-making group for small community grants. Grant-making criteria, desired outcomes and evaluation criteria all had to be developed by the local community. The purpose of the developmental process was to support internal, community-based accountability (as opposed to external judgment by the affluent and distant board of the sponsoring foundation). My role, then, has been to support the local community in developing its own evaluation process and sense of self-accountability.

The process is in its fourth year. Things are turning out quite differently than anyone could have foreseen. There isn't space here to share those details, nor are they particularly important. The point is that the evaluation process has had to be highly flexible and responsive. Taking a 20-year, developmental perspective, where the locus of accountability is community-based rather than funder-based, changes all the usual parameters of evaluation.

4. *A Reflective Practice Process in Adult Education.* I've been working for several years with a suburban adult and community education program in facilitating a reflective practice process for staff development and organizational change. We meet monthly during which times we get reports from staff about their action research observations for the last month. The focus of these observations is whatever issue the group has chosen the previous month. The reflective practice process involves:

1) identifying an issue, interest, or concern;
2) agreeing to try something;
3) agreeing to observe what is tried;
4) reporting back to the group individually;
5) identifying patterns of experience or themes across the separate reports;
6) deciding what to try next, that is, determining the action implications of the findings; and
7) repeating the process with the new commitment to action.

Over several years this process has supported major curricular and organizational change. Evaluation is ongoing and feedback is immediate. The process combines staff and organizational development and evaluation. My role as facilitator is to keep them focused on data-based observations and to help them interpret and apply findings. There are no formal reports and no formative or summative judgments in the usual evaluation sense. There is only an ongoing developmental process of incremental change, informed by data and judgment, that has led to significant cumulative evolution of the entire program. They have become a learning organization.

ISSUES IN DEVELOPMENT EVALUATION

One reaction I've had from colleagues is that these examples aren't "evaluations" at all but rather organizational development efforts. I won't quarrel with that. There are sound arguments for defining evaluation narrowly in order to distinguish genuinely evaluative efforts from other kinds of organizational mucking around. But what I'm reporting is that, in each of the examples I've shared, and there are many others, my participation, identity, and role are considered evaluative by those with whom I'm engaged (and by whom I'm paid). There is no pretense of external independence. My role varies from being evaluation facilitator to full team member. In no case is my role *external* reporting and accountability. This may be a different role for an "evaluator," but I include "developmental evaluation" among the things we evaluators can do in part to open up the market for organizational development to evaluators. What we lose in conceptual clarity and purity with regard to our official definition of evaluation (whatever that may be), we gain in market share for evaluation expertise (essentially competing with organizational development types under the guise of "developmental evaluation").

When discussing the developmental evaluation option, I've heard concern from evaluators who emphasize the distinction between gathering data and rendering judgment versus offering recommendations and designing programs. Scriven (1994) has been articulate in arguing that being able to identify that something is or is not working (an evaluative function) is quite different from knowing how to fix or improve it (a design function). I agree that recommendations ought not be considered automatic. Indeed, in my own work, I prefer to facilitate the generation of recommendations by my clients and primary users. I rarely formulate independent recommendations. However, in the developmental evaluation process, part of my value to a design team is that I bring a reservoir of knowledge (based on 25 years of practice) about what kinds of things tend to work and where to anticipate problems. Young and novice evaluators are well-advised to stick close to the data. However, we gray heads have typically accumulated a great deal of knowledge and wisdom about what works and doesn't work. I think Scriven underestimates the valuable role evaluators can play in design and program improvement based on cumulative knowledge—the gray head effect.

Again, I'm looking at this in part from the perspective of marketing and positioning the profession. We know a lot about patterns of effectiveness, I think—and will know more over time. That knowledge makes us valuable partners in the design process. Crossing that line, however, does reduce independence of judgment. The costs and benefits of such a role change must be openly acknowledged and carefully assessed. For my part, I like blurring the lines and use the term "developmental evaluation" to legitimize doing so. It also helps clients to have a name for the process.

A FORMAL DEFINITION

In hope of validating the designation "developmental evaluation" as an evaluation option, I submitted the following definition to Scriven as a candidate for his fifth edition of *Evaluation Thesaurus*. In the event it doesn't make his cut, readers can reference it here.

> *Developmental evaluation*: Evaluation processes and activities that support program, project, product, personnel and/or organizational development (usually the latter). The evaluator is part of a team whose members collaborate to conceptualize, design, and test new approaches in a long-term, on-going process of continuous improvement, adaptation, and intentional change. The evaluator's primary function in the team is to elucidate team discussions with evaluative data and logic, and to facilitate data-based decision-making in the developmental process.

One final note of caution. The term "development" carries negative connotations in some settings. Miller, in *The Book of Jargon*, defines development as "a vague term used to euphemize large periods of time in which nothing happens" (1981, p. 208). Evaluators are well-advised to be attentive to what specific words mean in a particular context—and to choose their labels accordingly.

UTILIZATION-FOCUSED EVALUATION

When I have discussed developmental evaluation with colleagues and students they typically ask how it relates to utilization-focused evaluation (Patton, 1986). The short and simple answer is that developmental evaluation is one option in a utilization-focused process. But taking up the question gives me an opportunity to clarify where I think utilization-focused evaluation fits in the grander scheme of things and correspondingly to take issue with where it is typically placed in schemes of evaluation. First, a definition:

> *Utilization-Focused Evaluation is a process for making decisions about and focusing an evaluation on intended use by intended users.*

Utilization-focused evaluation shifts attention from methods or the object of evaluation (e.g., a program) to the intended users of evaluative processes and information, and *their* intended uses. The evaluator, rather than acting as an independent judge, becomes a facilitator of evaluative decision-making by intended users. Thus, the utilization-focused evaluator facilitates judgments about merit and worth by intended users rather than highlighting his or her own judgments. This approach recognizes that no evaluation can be value-free, therefore the values that inform the evaluation should be those of primary intended users. Utilization-focused evaluation is highly personalistic and situational. It is subject to external credibility attacks in that the evaluation facilitator develops a close working relationship with intended users. Therefore, trade-offs between independence and use should be addressed explicitly in the negotiating process. Where intended users want and need an independent, summative evaluation, such can be designed. Any kind of purpose (formative, summative, developmental), any kind of data (quantitative, qualitative, mixed), any kind of design (naturalistic, experimental) and any kind of focus (process, outcomes, costs, cost-benefit, etc.) can result. Choices among the many alternatives in evaluation are driven by attention to utility.

In typologies of evaluation, utilization-focused evaluation is typically considered a decision-oriented model. From my perspective, that's an inappropriate placement. Decision-oriented models focus the evaluation on major decisions that must be made about a program. That can occur in a utilization-focused process, but it is not automatically the focus. The focus is on the utility needs of specific people. Those needs may or may not include making major decisions about the program. Utilization-focused evaluation

should be placed in a category of user-oriented approaches, not decision-oriented approaches. The distinction for me is that this approach focuses on people (*intended users*). Sometimes this involves decisions they must make. More often, it involves incremental improvement processes, conceptual insights, and reality-testing their assumptions or beliefs. Increasingly, in my practice, it involves participating as a member of a development team.

A second common confusion is associating utilization-focused evaluation with qualitative methods. That confusion arises, I assume, from the fact that I've written about qualitative methods in evaluation (Patton, 1990). But utilization-focused evaluation is inherently multi-method and must be so if intended users are to have a wide variety of methodological options.

A third objection I hear commonly is that the term "utilization" is too academic. Many have suggested "use" instead of utilization. I value simplicity and have no problem with the shorter word, though I admit that, to my ear, "utilization" connotes a process whereas "use" connotes a more defined action. Because utilization-focused evaluation is a process that considers multiple uses, I've opted to maintain the designation in my writings. Were I to rename the approach, it would not be use-focused evaluation, but user-focused evaluation.

DEVELOPMENT AND USE

The notion of developmental evaluation really calls into question three traditional mainstays of evaluation: purpose, evaluand, and timeline. These three are intertwined in the classic definition of program evaluation as determining whether the program's goals have been met. This definition assumes a fixed program, a delimited time period, and a goal-attainment purpose. Developmental evaluation is a way of being useful in innovative settings where goals are emergent and changing rather than predetermined and fixed, time periods are fluid and forward-looking rather than artificially imposed by external deadlines, and purpose is learning, innovation, and change rather than external accountability (summative evaluation) or getting ready for external accountability (formative evaluation).

Developmental evaluation has emerged as primary in my own practice because it affords unusual opportunities for me to be useful in working on issues I care about with imaginative and committed people whose values I respect. At this stage in my practice and life, after many years of project evaluation, developmental engagement on important issues matters a great deal to me.

REFERENCES

House, E. R. (1993). *Professional evaluation.* Newbury Park, CA: Sage.
Miller, D. E. (1981). *The book of jargon.* New York: Macmillan.
Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed). Newbury Park, CA: Sage.
Patton, M. Q. (1986). *Utilization-focused evaluation* (2nd ed). Newbury Park, CA: Sage.
Rheingold, H. (1988). *They have a word for it.* New York: St.Martin's Press

Scriven, M. (1994). *Evaluation thesaurus* (5th ed.) Newbury Park, CA: Sage.

Stockdill, S. H., Duhon-Sells, R. M., Olson, R. A., & Patton, M. Q. (1992). Voices in the design and evaluation of a multicultural education program: a developmental approach. In A. M. Madison (Ed.), *Minority issues in program evaluation. New Directions in Program Evaluation, 53*, 17-34.