

Papers in this section focus on evaluation from a historical perspective. They may analyze important turning points within the profession, provide commentary on historically significant evaluation works, or describe and analyze what promises to be a contemporary watershed event with important implications for the future of evaluation. If you have any questions or suggestions about topics you would like to see addressed in this section or would like to chat about an idea you are considering for submission, feel free to e-mail John Gargani at jgargani@berkeley.edu.

The Oral History of Evaluation, Part 5

An Interview With Michael Quinn Patton

The Oral History Project Team

The Oral History Project Team is interviewing people whose signal contributions to program evaluation have shaped the evolution of the field over time. Our goal is to capture the professional evolution of those who have contributed to the way evaluation in the United States is understood and practiced today. The following interview, conducted by Jean King and Lija Greenseid, presents the evolution of Michael Quinn Patton.

Michael Patton is an evaluation and organizational development consultant who spent 18 years at the University of Minnesota (1973-1991), including 5 years as the director of the Minnesota Center for Social Research and 10 years with the Minnesota Extension Service. His writings and practice have influenced the field in numerous key areas. The publication of his book *Utilization-Focused Evaluation* in 1978 focused attention on the importance of intended users and uses of evaluation information and, in its third revision (Patton, 1997), of the evaluation process. His book *Qualitative Research and Evaluation Methods*, now in its third edition (Patton, 2002), was one of the first qualitative methods texts. Among Michael's awards are the Alva and Gunnar Myrdal Award and the Paul F. Lazarsfeld Award from the American Evaluation Association (AEA; acknowledging his contributions to both practice and theory), the Lester F. Ward Award for Outstanding Contributions to Applied Sociology from the Society for Applied Sociology, and the University of Minnesota's Morse Amoco Award for outstanding teaching.

Interview With Michael Patton

Lija: In the 1970s, you conducted a seminal study on the use of evaluation that became the basis of your book *Utilization-Focused Evaluation*. Why did you conduct that study?

Authors' Note: The Oral History Project Team includes Jean King, Mel Mark, and Robin Miller. Special thanks to Lija Greenseid, University of Minnesota, for her assistance with this interview.

American Journal of Evaluation, Vol. 28 No. 1, March 2007 102-114
DOI: 10.1177/1098214006298129
© 2007 American Evaluation Association

Michael: That study was the centerpiece of a shared experience we designed for participants in the University of Minnesota's Evaluation Methodology Training Program. I was the first postdoctoral fellow in the National Institute of Mental Health (NIMH) Evaluation Methodology Training Program that John Brandl, distinguished director of the School of Public Affairs, brought to the University of Minnesota. As I remember it, five universities were awarded grants when NIMH decided to support interdisciplinary evaluation methodology training. Some 17 different departments collaborated on the grant at the University of Minnesota. The first year of that program was focused on applying sophisticated methods to evaluation. The dominant notion at the time was that the way to get more use was to do more sophisticated studies. One of the first experiments we looked at was RAND's evaluation of the Alum Rock educational voucher demonstration, the first attempt to install vouchers within a public school system.

I became director of the training program the second year of the grant, 1974, and wanted to be sure that predoctoral and postdoctoral fellows were actually doing evaluations rather than just academically studying evaluations done by others. So participants in the program began going out to schools and small agencies to assist with evaluations and found that staff were having trouble making sense of percentages, much less dealing with standardized regression coefficients and path analysis models. As I said, the evaluation methodology training program was originally designed to teach sophisticated evaluation models and research methods. But as program participants undertook real evaluations in local settings, we found much of our traditional methodological training to be irrelevant. We learned that evaluators need skills in building relationships, facilitating groups, managing conflict, walking political tight ropes, and effective interpersonal communications. Technical skills and social science knowledge weren't sufficient to get evaluations used. People skills were also critical. Evaluators without the savvy and skills to deal with people and politics will find their work largely ignored or, worse yet, used inappropriately.

We learned that a particular evaluation may have multiple levels of stakeholders and therefore need multiple levels of stakeholder involvement. We learned that the sophisticated methodological techniques that were highly valued for dissertation research had little applicability for small-scale, local, formative evaluations. We had to develop methods, including mixed methods and qualitative approaches, which were appropriate and responsive to local needs.

The disjuncture between the original, purely methodological focus of the training program and local evaluation needs was such that we decided we would make evaluation use the centerpiece of the program to provide focus and coherence. I think we awarded six graduate fellowships and two postdocs each year, and they took a year-long seminar together. Enhancing use became the core theme of the seminar. How do you design evaluations for use? How do you implement for use? What kinds of reporting contribute to use? The seminar gave rise to the idea of conducting our own utilization study of federal health agencies—and those findings led to my book *Utilization-Focused Evaluation*.

Attention to use derived from my values orientation. After a stint in the Peace Corps in the 1960s, I went into a new Sociology of Economic Development program, which was also a NIMH program, at the University of Wisconsin, Madison. I had an applied orientation to sociology, wanted to do work that was useful, and happened upon evaluation as a dissertation topic. That led me to the University of Minnesota as a postdoc in evaluation methodology. So I was very much personally and philosophically oriented toward use, and studying use was a way to try to provide coherence to the evaluation methodology training program.

Lija: Would you evaluate the NIMH program? Did it succeed? To what extent and in what ways did the program shape your professional identity?

Michael: Did it succeed? That's a very political, values-based question. Of course, judging success depends on what criteria one applies. So let me describe the alternative criteria that were in play. First, there's the question of who benefited from participation in the program. The program involved more than just the people who were formally getting their doctorates. It led to my doing the year-long evaluation seminar, which attracted students from across the university interested in or engaged with evaluations. I had 20 students in the seminar each year, fewer than half of whom were on program fellowships. The people who participated in that seminar went on to populate most of the major evaluation units in Minnesota. For example, the superb Ramsey County evaluation unit had

several people participate, including Gene Lyle, an AEA Myrdal Award recipient. The nationally recognized Hazelden Foundation chemical dependency program had evaluation staff take the seminar. Leading local evaluation and research units sent staff to the seminar as did state government agencies and staff in the not-for-profit sector. So over the 5 years that I convened that seminar as the centerpiece of the evaluation methodology training program, a substantial number of participants had positions or got jobs in influential organizations actually doing evaluations. People who came through that seminar became staff to important state legislative committees. A number of seminar participants joined the well-respected evaluation unit of Minnesota's legislative audit agency.

That seminar also led to my starting to do evaluation workshops through the university's continuing education program. We were both creating and fulfilling a demand for evaluation training. Those 1-day evaluation workshops educated a large number of evaluation consumers. Indeed, I've told people in the workshops I do on how to build an evaluation consulting practice that the purpose of a 1-day evaluation workshop open to all-comers is not to train people how to do evaluation, but to train consumers. You can't train people to actually do an evaluation in a day, but you can bring them to understand what evaluation is, why it is useful, why they need to support evaluation, and how to find a competent evaluator. So those workshops, which I did four times a year for 5 years, created demand. Indeed, I still get an occasional phone call from somebody who took one of those workshops in the mid-70s. Maybe they're moving offices or cleaning out files and come across the workshop handouts; it turns out that they're finally getting around to doing an evaluation, and they contact me and ask if I'm still doing evaluations. As I mentioned, I now do a consulting workshop as part of The Evaluators' Institute, and that's one of the things I emphasize in that consulting workshop—that we have to create demand for useful evaluations, and one of the ways to do that is to train consumers.

So deciding to make evaluation use the central theme of the evaluation methodology training program led us to the research on use that became utilization-focused evaluation. Those early continuing education workshops and our research on use led to my doing workshops for the Evaluation Network annual conferences beginning in 1975, workshops I've done every year since, but now for the AEA. So, if the evaluation methodology training program is judged by its impact on the Minnesota evaluation scene—placing evaluators in local and state government units doing useful evaluations, staffing evaluation in not-for-profit agencies and philanthropic foundations, building a model of utilization-focused evaluation, and creating demand for evaluation—then by those outcomes, I think the program can be considered quite successful. It was also a success in expanding attention to evaluation within the university. We had faculty from urban geography, a strong department at the University of Minnesota. We had faculty from statistics, education, psychology, political science, economics—virtually all the social sciences. We had faculty from the law school participate, people from public health and medical departments, and staff from agricultural extension. It was a truly interdisciplinary program.

But there are other criteria and another side to the question of whether the program was successful. When it came up for renewal by NIMH, the external review team they sent consisted of two social science deans from major universities, the head of a well-respected sociology department, and the director of the program from NIMH. None of them, by the way, had any expertise in evaluation. Their criteria were whether we were teaching what they considered a national curriculum (based on comparison to the other four funded training programs), whether we were placing graduates of the program in tenure-track positions in major universities and national research organizations, and whether we were teaching experimental designs, sophisticated statistics, and cost-benefit analysis as the core methods of the program. They considered the program's emphasis on mixed methods a weakness. They considered our qualitative study of use in federal agencies to be no more than a weak seminar project and not serious scholarship. The fact that we had established ourselves as the premier training ground for local evaluators, a fact they did not dispute, was judged a negative because it gave the program a professional rather than scholarly flavor. They viewed evaluation as an emerging subdiscipline of social science. They wanted the program to produce evaluation research scholars. We had focused the program on training professional evaluation practitioners who could conduct useful evaluations. They didn't think that local evaluation was important. As a federal agency, the NIMH director wanted us to be conducting evaluations of federal programs. That's what

they considered the purpose of the program. All in all, the official site visit evaluation was extremely negative and the program was not refunded.

By the way, it was during the site visit that I heard these NIMH criteria for the first time. I was completely blind-sided. I was very proud of the program and looked forward to the opportunity to show off what we had accomplished. I could not have been more surprised. I learned a lot from that experience, especially what happens when there is no communication between a funder and grant recipient about evaluation criteria—we never received any feedback on our annual reports—and then what it's like to be evaluated by independent, external evaluators who bring and impose their own criteria from the funder and have no interest in what has emerged at the ground level in practice.

Now from a utilization perspective, the summative evaluation by the external evaluation team was certainly used. Indeed, it was the primary basis for the nonrenewal decision. As program director for 4 years, I had no idea how the program was going to be evaluated or even that it was going to be evaluated. The stakes were high and, unbeknownst to me, it turned out I wasn't viewed as a stakeholder. My criteria of success didn't matter, nor did the criteria of the program's faculty advisory committee, and I was in no way involved in negotiating the external evaluation focus, design, questions, or use. It was a huge learning experience.

When the NIMH program was terminated, the university lost interest in an interdisciplinary approach to evaluation. The program had no home and no funding. Our refunding proposal envisioned making the NIMH Evaluation Methodology Training Program the foundation of an interdisciplinary Evaluation Institute that would have as its mission integrating theory, research, and practice across departments and colleges to support the emerging evaluation profession. That was in 1978, mind you. Had our proposal been supported, the University of Minnesota would have had something like the Center for the Study of Evaluation that Marv Alkin created at UCLA or the Evaluation Center that Dan Stufflebeam created at Western Michigan. But those important places were grounded in educational evaluation. Our vision was to create a truly interdisciplinary evaluation institute. To this day, no such place exists. I went on to become a solo practitioner, alienated from both the university and the federal bureaucracy, though I realize after all these years that both were doing what they were designed to do. Their respective missions, as their leadership interpreted them, just didn't match what I wanted to do and my vision for evaluation. *C'est la vie.*

Lija: In addition to writing and publishing in evaluation, you're a well-known writer on qualitative methods, and yet your early training was in the hardcore, social science, quantitative methodologies at the University of Wisconsin. To what extent have your epistemological beliefs changed over time?

Michael: I do 30 or 40 workshops a year on evaluation or qualitative methods or some combination of those. I often begin by acknowledging that I've never had a course in either evaluation research or qualitative methods. All my doctoral training was in survey research, experimental designs, and statistical analysis. There was no qualitative course offered in sociology at the University of Wisconsin, Madison, when I was there. Places like Chicago and Northwestern had qualitative courses in sociology, but Wisconsin was known as a quantitative department.

I got into qualitative methods in my dissertation because I had an opportunity to evaluate a statewide open education program in North Dakota and did my dissertation on open education as a form of organizational innovation. The open education people there were quite hostile to standardized testing, to applying numbers to their kids. But I took on the evaluation because my wife at that time was in the master's program at the University of North Dakota, and that gave me a way to be out there more often. The program was supported by federal funds as part of a national Trainers of Teacher Trainers initiative. North Dakota was one of about 35 sites that had that money. My dissertation support was from the evaluation budget of the program. So I began negotiating the evaluation with them. I agreed to do the study before I understood what it was going to involve. At the first meeting after I had accepted the project, they told me that they were very excited to have me do my dissertation on their program, then they said, "We know this is your dissertation and you're going to have to do what you need to do to get a dissertation out of this. We're glad to have you do this. You can do anything you want to do as long as you don't use any numbers. We don't want our kids labeled and stuck into numerical categories." And I looked at them with what must have been shock on my face and said, "You're joking, right?"

They said, “No, we don’t want to fold, spindle, or mutilate our kids. We’re not going to put them in categories, but you can do anything else.” And I said, “What else is there?” And thus began my education in qualitative inquiry, largely experientially learned, with collaboration from some excellent people along the way.

People involved in open, humanistic education from across the country were struggling with evaluation. They came together at the invitation of Vito Perrone, dean of the Center for Teaching and Learning at the University of North Dakota. The group dubbed itself the North Dakota Study Group on Evaluation. Vito Perrone and North Dakota Study Group participants taught me about qualitative methods, conducting case studies, interviewing people, and doing classroom observations. I dedicated the first edition of *Utilization-Focused Evaluation* to Vito. To conduct the evaluation of the open education program, I had a team of three colleagues who did all the interviewing and observational work with me. But I had to do regression analysis to get a doctorate at Madison. So we had to code all our qualitative classroom observations on quantitative scales, a set of organizational innovation dimensions, so that I could run a regression analysis looking at the relationship between dimensions of innovativeness and educational outcomes. The people in North Dakota got the qualitative evaluation but never saw the quantitative analysis, and the faculty on my doctoral committee at Wisconsin never saw the full qualitative study, just the statistical analysis. And that’s how I learned firsthand about the paradigms debate. It wasn’t an abstract, intellectual argument for me. It had direct implications for what I had to do in working with people on different sides of that debate.

The story of how I came to write the *Qualitative Research and Evaluation Methods* book is also something of a saga. Sara Miller McCune founded Sage Publications with her husband George. *Sage* derives from the first two letters of their names, Sara and George. In 1977, when I had completed the first five chapters of *Utilization-Focused Evaluation*, I went to the American Sociological Association meeting in Chicago with a prospectus and draft chapters and dropped them off at a number of booths of book publishers, one of which was Sage, which had published the *Handbook of Evaluation Research* in 1975, but had only a few books out at that time. Within an hour, I got a call from Sara McCune. She had read my materials, understood the proposed book, was conversant about its niche in the emerging field of evaluation, and had a marketing plan in mind. I was leery, not knowing Sage. I asked for time to consider and then checked back with other publishers where I had dropped off my prospectus. The people at those booths were all salespeople, and all they could tell me was that they would pass my prospectus on to an editor who would send it out for review, and I shouldn’t expect to hear anything for 3 to 6 months. Sara McCune, in contrast, offered me a contract and had a plan in mind about how Sage would promote the book. I did a quick evaluation of the alternatives and made a summative decision to sign with Sage.

Now, bear with me, I’m getting to the story of how I came to write the qualitative book. In *Utilization-Focused Evaluation*, I included a chapter on the qualitative/quantitative paradigms debate. It was based on a monograph entitled *Alternative Evaluation Research Paradigms* I had written for the North Dakota Study Group on Evaluation. I adapted that monograph for a chapter in *Utilization-Focused Evaluation*.

Not long after Sage published *Utilization-Focused Evaluation*, Sara McCune, knowing of my interest in the paradigms debate from that chapter in the book, sent me a qualitative research manuscript to review. She said that she wanted Sage to do a book on qualitative evaluation. Qualitative methods were getting a lot of attention, and she wanted to get a book out on qualitative evaluation ASAP. I read the manuscript immediately and called her back with my reaction, which was something like, “Sara, this manuscript is horrible. This is not about qualitative methods. It’s a book about how to do investigations. It’s an auditing site visit book that assumes people are lying to you. It’s written from the perspective of how to uncover problems in field visits to programs. You can’t publish this as a basic qualitative evaluation book. This will do enormous harm to the field. This isn’t a balanced presentation of either qualitative methods or program evaluation. It’s a very small part of one aspect of the field.” She heard me out and replied that Sage really needed a book on qualitative evaluation and didn’t have any other manuscript in hand and needed to get something out quickly to capture the market. She asked me to write the book.

I had a toddler son and a newborn and was very caught up in parenting. I had only recently finished *Utilization-Focused Evaluation* and was still in recovery from that. I was directing the Minnesota Center for Social Research at the University of Minnesota and had a lot of projects to do. So I declined. I told Sara that I didn't know enough to write a qualitative book, that I had no formal training.

She said she wanted to come out and talk with me about it. She flew to Saint Paul from San Francisco, had dinner with us, and convinced my wife that this book was needed. Sara cajoled, charmed, and persisted. At the end of the evening, she told me the bottom line is: If you don't write a book, we're going to have to publish the manuscript you saw. So, I rearranged my life and went into overload and churned out the qualitative book.

I never intended to develop expertise in qualitative evaluation and didn't have very many good cases to include in the first edition. But that book came out in 1980 and was their first qualitative evaluation book. They now have what seems like scores of them, and *Qualitative Research and Evaluation Methods* is in its third edition. The fourth edition of *Utilization-Focused Evaluation* will be out next year. Sage published two other evaluation books I did, *Practical Evaluation* (Patton, 1982) and *Creative Evaluation* (Patton, 1981). My long and fruitful association with Sara McCune and Sage Publications has coincided with Sage becoming the world's premier publisher of evaluation books and qualitative methods books. Some years ago, AEA gave Sara Miller McCune a special award recognizing Sage's contributions to the development of the field of evaluation. Last year, Sage celebrated its 40th anniversary, and I flew to New York City to be part of the celebration. Sage remains one of only a few independent publishers. I'm not doing public relations for Sage Publications. I'm just wanting to recognize that the long-term association with Sage has been a central storyline in my evaluation career.

Lija: You have stayed in Minnesota all these years. Why did you decide to remain here?

Michael: As much as anything, it had to do with Minnesota being a great place to raise a family. I had just become a father when I first moved to Minnesota after graduate school. I got to know the community well during the years I directed the Evaluation Methodology Training Program and then as director of the Minnesota Center for Social Research at the University of Minnesota. By 1980, at any one time we were doing 30 to 40 small evaluations for state programs, local agencies, and foundations. This provided the foundation for a solid consulting career and meant I didn't need to worry about tenure. You actually can't just pick up and start a consulting business someplace else, especially if you are doing local-level evaluation, which is what I was specializing in, where I could get close enough to the primary intended users to really work with them directly and interpersonally. I had projects with all the major philanthropic foundations in Minnesota and helped introduce them to evaluation. I worked with the leading nonprofits and with state government agencies. The kind of local-level, utilization-focused evaluations I did made building trusting relationships important.

At the same time, I found little support for or appreciation of evaluation at the University of Minnesota. Having come to the university as a postdoctoral fellow—not on tenure track—and having strayed from sociology, my doctoral field, I didn't fit at the university. I found a home for a while doing international cooperative extension programming, but that was a marginal, soft money position. I never applied for a tenure track university position in Minnesota or elsewhere. I had seen enough of university politics to know it would be a diversion from what I wanted to do. And it seemed highly unlikely that any department would have ever granted me tenure for the type of work I did. So I decided not to go that route, which meant that having built a base in Minnesota where I could practice my craft, this was the place to stay.

The downside of running a soft money center at the university was that by 1980, I had some 15 people dependent upon me to continue to find projects and raise funds to support them. The evaluation work we did at the Minnesota Center for Social Research supported a number of graduate students as well as regular professional staff. And I was the principal investigator on all of those projects. I was constantly writing proposals, meeting with clients, and overseeing the work. It was 16 hours a day. It was very exciting work. I was young and energetic, but I was also realizing that the harder I worked and the more successful we became, the more responsibility I had for the work of others. I also wanted to do more writing. The only way to get off the treadmill and break free was to leave the country.

Right at that time, the Big Ten universities got a large USAID (U.S. Agency for International Development) project doing planning and evaluation in the Caribbean through the Minnesota Extension Service, and I applied to be the project director of that initiative. Because I had some cooperative extension background, had done my master's in rural sociology, and had been an agricultural extension agent in the Peace Corps in Burkina Faso, I got serious consideration for the job. What tipped the scales in my favor was the evaluation expertise I would bring. USAID was getting really big on evaluation at that time, so I got the job. Sadly, I left Minnesota winters to work and live with my family in Trinidad for a couple of years. That whetted my appetite for international evaluation work, which has been a major focus since that time.

I returned to Minnesota in 1982 and had to reestablish my consulting practice, but I did that without creating a center. By then, I didn't want to run an organization. I didn't want to supervise people. I continued part-time with the Caribbean project for 10 years. Throughout the 1980s, I directed that project, which gave me a half-time base. The rest of my time I devoted to consulting and writing to build my evaluation practice back up again. It all seems like a long time ago.

Jean: In an earlier AEA oral history interview, Will Shadish mentioned that you and he once had a conversation about your identities as evaluators. He stated that you "definitely" felt like an evaluator, where he was less comfortable with the title. Did he get it right?

Michael: He did get that right. Indeed, a piece of that history is that when I was president of AEA in 1988 and we were reviewing various organizational documents and the constitution, I noted that on the application to become a member of AEA, being an evaluator was not one of the options. You could only designate yourself as identifying with a discipline. You could be a sociologist, psychologist, economist, or an educator, but you could not be an evaluator. So we instituted the change in the membership application so that identifying oneself as an evaluator was an option.

Jean: When did you first embrace the identity of evaluator?

Michael: It probably happened in conjunction with my becoming the director of the Evaluation Methodology Training Program in '74 and then director of the Minnesota Center for Social Research (MCSR) at the University of Minnesota in '75. MCSR became an institutional setting for doing evaluations, mostly small-scale evaluations, primarily in the state and mostly in the Twin Cities. I was a new PhD in sociology, which had no caché with anybody and scared potential clients to death; so, as bad as it was to be an evaluator, it was a whole lot more friendly, concrete, and understandable than being a sociologist. So I started calling myself an evaluator then. We were doing evaluations, we were bidding on evaluations, and it was actually easier to explain what an evaluator did than what a sociologist did. Still is.

Jean: So you're not an "accidental evaluator"?

Michael: Not at all. I gave up the sociology identity within a year of graduating with my PhD and never looked back. I should add that while I quickly took on an identity as an evaluator rather than a sociologist, my sociological roots have had a sustained influence on how I view the world and have given me many of the core concepts I use to make sense of situations. My graduate studies in the sociology of knowledge, the nature of power and conflict, diffusion of innovations, organizational sociology, and sociological theory and methods are enduring influences. However, in coming to appreciate the importance of the personal factor in explaining evaluation use, I had to unlearn a lot of sociology. The dominant Weberian perspective in organizational sociology posits that organizations are made up of and operate based on positions, roles, and norms such that the individuality of people matters little because individuals are socialized to occupy specific roles and positions and behave according to specific learned norms, all for the greater good of the organization's goal attainment. In studying use and engaging in evaluation practice, I've found that individual people make a huge difference. It's not just about structures, positions, and roles. People matter. Individual people matter a lot. And that's not the central message of or basic wisdom derived from sociology.

Jean: In what ways would you say over the 30 years since that time your identity as an evaluator has changed?

Michael: A couple of changes have occurred. The first is a function of the way I most often work these days, but emerged formally as a result of a conversation with Michael Scriven about our respective approaches to evaluation and other conversations over the years about the roles that we, as evaluators, play. The most widely used definition of evaluation emphasizes judging the

merit, worth, and significance of something. By that definition, program evaluators have the job of determining the merit, worth, and significance of programs and making independent judgments about effectiveness and efficiency. However, when I facilitate a utilization-focused evaluation, I work with primary intended users to arrive at informed judgments about the program in question. I rarely make the judgments myself, alone. I facilitate other people making judgments. Scriven insisted that when operating this way, I wasn't an evaluator at all. So I've come to call myself an "evaluation facilitator." I facilitate, coach, train, and otherwise build the capacity of people to do evaluations. Sometimes I still take on a contract where my independent judgment is sought and I play the role of evaluator, but that's not my most common practice.

Now, with increased understanding of and attention to *process use*, where people learn not just from the findings of an evaluation but also from participating in the evaluation process and learning to think evaluatively, I often describe myself as a facilitator of *evaluative thinking*. I facilitate not just specific evaluations but facilitate building a culture of evaluation in organizations and training leadership to engage evaluation findings and use evaluative thinking in all aspects of their work.

Jean: Do you think you're unique in the field in that?

Michael: I'm probably more intentional and deliberate about it than most others. My work crosses and integrates organizational development and evaluation. It's actually quite a substantial niche where I'm finding lots of demand. Plus, working this way has helped keep me out of the business of writing reports, which I never much cared for anyway, and which I generally try to avoid. I occasionally do a report when I take on a contract to actually do an evaluation, but I've mostly been out of that work for about 10 years.

Jean: Do you consider yourself an evaluation researcher?

Michael: That's a complicated question. Let me distinguish three different senses of what it means to be an evaluation researcher as least as I see it. One meaning, which derives from the former Evaluation Research Society, defines evaluation as applying social science methods to the study of program effectiveness. That's the tradition of Rossi and Freeman. From my perspective, theory-based evaluation is a direct descendant of this tradition, and much theory-based evaluation research is more interested in testing social science theory in program settings than in determining the effectiveness of the particular program being evaluated. In this use of the label "evaluation research," the emphasis is on *research*, and social science researchers are pursuing their specific research interests in an evaluation context or with evaluation funds. I began in that tradition because initially, when I was doing my dissertation, I simply saw evaluation as a way to pursue my more general scholarly interest in organizational sociology, diffusion of innovations, and sociology of knowledge with an emphasis on utilization. As I noted earlier, I quickly became fascinated with evaluation as its own field and abandoned the more general social science interest. I'm no longer an evaluation researcher in that original meaning of the phrase, but a good many people are, and that remains one way in which people engage in evaluation, as evaluation researchers—emphasis on *research*.

A second meaning of evaluation researcher refers to those of us who do research *on* evaluation. In that sense, I am most definitely an evaluation researcher. I've participated in many such studies, including the very first utilization study of federal health evaluations I did that was the basis for *Utilization-Focused Evaluation*. I was a part of the study that the AEA Topical Interest Group (TIG) on evaluation use did, implemented by Hallie Preskill, Lynn Shula, and Valerie Caracelli (Preskill & Caracelli, 1997). I advised on design of that study, was at the TIG meeting where we talked about doing that, reviewed the instrument, and participated in interpreting the results. I advise colleagues on studies of evaluation use as, for example, in the case of a major study of use being supported by NSF (National Science Foundation) on which one of the principal investigators is my good friend and colleague, Jean King, at the University of Minnesota. The AEA Oral History Project Team is also gathering case stories from ancient, worn-out, over-the-hill, one-foot-out-the-door-but-somehow-still-alive-and-kicking evaluators. I would consider that a form of evaluation research, as was Marv Alkin's book on *Evaluation Roots*, which also captures such stories, though in less detail.

In that tradition of evaluation research, then, I've regularly been a part of sessions at AEA discussing and helping people design utilization studies. I've been active in the TIG on theory, which

has included examining theories about the nature of evaluation. I am an evaluation researcher in evaluating my own practice. I've followed up and have long made it a practice to follow up evaluations I do—or facilitate—to find out how they are used. It was that follow-up work that led to the conceptualization of process use in the third edition of *Utilization-Focused Evaluation*. So my subsequent research on evaluation became evaluation of my own evaluations. In that sense, I'm an evaluation researcher—one who does research on evaluations.

A third, newer meaning of evaluation researcher involves the generation of generic knowledge about patterns of program effectiveness. Most evaluations determine the effectiveness of a specific program. Formative and summative evaluations are aimed at particular programs. But increasingly, we have occasion to look across a number of programs and their separate evaluations in search of generalizable lessons and generically effective practices. This is a knowledge-generating use of evaluation, and those involved in identifying such lessons learned or conducting meta-analyses are evaluation researchers. This is an arena of great interest to funders and policy makers, and I've done some of that kind of evaluation research.

Lija: You mentioned that you discussed the qualitative-quantitative debate in your first edition of *Utilization-Focused Evaluation*. In the most recent edition, you talk about the happy resolution of that debate. As we know, that debate has resurged and the renewed paradigm wars are affecting all of us. How has this debate affected the field of evaluation?

Michael: Well, the debate has come and gone and now come again, and it changes in every iteration. The current debate is different in some important ways. The debate now is actually quite narrow. It's quite important, but it's actually quite narrow. It's about what constitutes credible "scientific" evidence of program impact in which the intervention can be empirically causally linked to measured outcomes. The earlier qualitative-quantitative debate was about credible evidence more generally at a time we didn't have as much sophistication about different kinds of evaluation, varying purposes, and a variety of methods matched to distinct purposes. We didn't have the history of distinguishing much among formative, summative, and developmental purposes, and many of the different models had not yet appeared or attracted attention, approaches like participatory, democratic deliberative, theory-driven, realist, and a host of other evaluation distinctions. We had not yet conceptualized the difference between findings use and process use.

Qualitative evaluation and qualitative methods in evaluation are actually not in dispute in the current debate. Indeed, it is recognized among the people who are pushing randomized control trials that you need implementation data and process data to know what's in the black box of experimental designs. So there has been an acceptance of an important niche for qualitative evaluation. The classic qualitative-quantitative debate was more about measurement and the relative value and meaningfulness of numbers versus narratives. That debate has been largely settled in favor of mixed methods and methodological appropriateness, focusing on matching purpose and method.

The current debate is more narrowly focused on how to determine causal attribution in summative impact evaluations, at least that's how I've experienced the debate as I've been involved in this contentious round of claims and counter-claims. This debate certainly touches on what you can do with case studies versus randomized control trials, but the focus is on a quite narrow paradigm of causality. The debate is about what kind of evidence and what degree of certainty is needed to be able to attribute the outcomes of a program to the intervention of the program. And, it seems to me, it's typically more about policy evaluation than about program evaluation. Unfortunately, the amounts of money at stake are huge, and the misunderstandings in the public mind are huge, so I find myself on the debate circuit again. I've been debating evaluation design alternatives at the World Bank, at NIH (National Institutes of Health) and NIMH, and in various international conferences. I'm getting more and more invitations to debate.

And the debate has changed from the early days of the qualitative-quantitative debate. The issue now, as I see it, is whether randomized experiments are *the gold standard* of design, that is, whether there is one best and preferred method, even for something as focused as impact evaluations. The Campbell Collaborative strongly advocates the gold standard position, and the leadership of the Campbell and Cochrane groups have great influence. Now I dispute that there is

a gold standard or should be a gold standard, and I support the AEA position against a gold standard, but it's important to recognize that this iteration of the debate is not calling into question qualitative evaluation in general. The problem with a gold standard is that it creates enormous and perverse distortions in how evaluation resources are allocated, valued, promoted, and mandated. It means that method determines the evaluation question because in order to design evaluations that meet the gold standard, evaluators are prematurely and inappropriately designing experiments just to meet the so-called gold standard, not because the program is ready for or would benefit from such a design. For example, I'm seeing experimental designs when the intervention still needs formative development before it is summatively evaluated. This is especially unfortunate internationally where I see many, many inappropriate designs as international funding agencies try to mandate adherence to the gold standard. It will take some time for these agencies to understand the useless and damaging nonsense that results, but nonsense it is. And so I predict that this overreaching by the adherents of the gold standard, while it will sadly waste a lot of resources, will die of its own weight because intended users will come to see that these designs don't deliver what is promised.

As the AEA position paper on this stated, there are lots of different kinds of evaluation and lots of different methods, including alternative ways of establishing causality and dealing with attribution. The randomized experiment true believers assert that there is only one scientific and credible way of establishing causality. They are wrong. The challenge is matching method to purpose and context, even for impact evaluations. The debate has very high stakes and will rage for a while, but I reiterate that I'm convinced it will die of its own weight because advocates of randomized control trials (RCTs) as the gold standard can't actually deliver on the claims they are making about what RCTs will deliver.

In the meantime, I'm actually enjoying being back in the debate. I've gotten to be a much better debater over the years. An added benefit is that I don't have many other places in my life to express anger, so given the many outrageous things going on in the world, it's helpful to have an outlet for anger, and that's become my outlet. I'm sorry to say that the hardcore gold standard advocates sometimes manifest disconcerting similarities with evangelical religious extremists who believe there is only *one true way*. Indeed, an OECD (Organization for Cooperation and Economic Development) handbook on evaluating micro-enterprise programs, written by Professor David Storey of Warwick Business School, posits "seven steps to heaven," where heaven is a randomized experiment. So, for the materially and worldly oriented, we have the advocates of the gold standard, and for the spiritually oriented, we now have the path to heaven, where heaven is an RCT. What they share with narrow-minded neo-con politicians and evangelicals of all persuasions is an intolerance for diversity and the belief that they have the one truth. The AEA position, and my own, envisions a world of methodological pluralism, dialogue, and honoring different perspectives. The gold standard folks are having none of that, so their leadership no longer comes to AEA national conferences. It's unfortunate—and will hopefully pass.

Jean: You mentioned your Peace Corps experiences early on, and you now have a jet-setting lifestyle, earning more frequent flyer miles than any other person I know. Can you name all of the countries that you have worked in?

Michael: We ought to distinguish between living and working in a country versus going for a few days to give a keynote speech and do workshops. It's the latter I do a lot of now. I lived and worked in Burkina Faso for 2 years. I lived in Tanzania for 3 months doing my master's research. I also lived in northern Peru for 3 months on an undergraduate intern. I lived and worked in the Caribbean for 2 years, headquartered at the University of the West Indies in Trinidad, but also spending a great deal of time in all of the Leeward and Windward islands from Antigua, St. Kitts and Nevis, Montserrat, and Barbados to St. Lucia, Dominica, Grenada, and St. Vincent the Grenadines. Belize in Central America was also part of that project. So those are places I've spent substantial time and lived.

As evaluation has become global, I've had the opportunity to travel the international circuit, helping launch some of the new national associations with inaugural keynote addresses and workshops. That usually involves being there between 1 week and 3 weeks, like Nairobi, Kenya, in

1999 for the launching of the African Evaluation Society and later the launching of the South African Evaluation Society. I've done some work in Japan with Masafumi Nagao, who translated *Utilization-Focused Evaluation* into Japanese and provided the leadership for formation of the Japanese Evaluation Society, so I was there for the launching of that organization. I've had the opportunity to visit most European countries for evaluation conferences, training, and consulting, as well as Australia and New Zealand several times. I'm in Canada a great deal, including working with Brad Cousins as cochair of the program for the international evaluation conference in Toronto in 2005. Leading up to that, I was in Brazil in 2004 for the launching of the Brazilian Evaluation Society, then back in Peru for the launching of the Latin American Evaluation Network in October of that year. I teach in the World Bank's International Program for Development Evaluation Training every year in Ottawa, which includes participants from throughout the world. So the dramatic growth in evaluation globally has provided wonderful opportunities to see the diversity of evaluation around the world.

International diversity is challenging our thinking about what constitutes good evaluation work and what it means for evaluation to be used in different cultural and political contexts. I'm currently revising *Utilization-Focused Evaluation* (4th edition), and one of the major areas for revision is adding more international examples and applications. And while I travel a great deal, I know a number of colleagues who travel even more. It comes with the territory when you're involved in training and supporting the growth of the profession as I have the privilege to do.

Jean: How do you think evaluation has changed as a result of this international boom?

Michael: It actually relates to the gold standard question because that narrow form of defining what "true" evaluation is, it seems to me, is challenged by the different cultural and political ways people think about knowledge, what constitutes knowledge, what constitutes evidence, how evidence impacts a political context, and the dramatically different role of nonprofits and governments in different places, to mention just a few areas of diversity and variability. I hadn't realized when I first went to Japan that there was no real not-for-profit sector, which was striking to me since most of my evaluation work had been in the not-for-profit sector. In a similar vein, I've been impressed by the ways in which different governments work, for example, in a parliamentary system, which is the model for much of the world. The Westminster model is much different than what we have here in the United States with our separation of the executive and legislative branches, and where most evaluation is an executive function. For evaluation to be a legislative function in many countries makes a difference in how evaluations are undertaken and reported. And, of course, there are the enormous cultural differences.

Sociologically, the United States has a very competitive and blaming culture. We like to find out who did something wrong, point a finger at them, and establish blame for shortcomings. We like to establish winners and losers. Japanese society, in contrast, greatly values social harmony. In Japan, the group takes responsibility, and for the sake of group harmony, they're not into blaming and embarrassing people and pointing out faults. I'm talking in broad generalizations here, I realize, but I've found perspectives on handling feedback and engaging in learning to be fundamentally different in Asian contexts. In Africa and Latin America, there are some strong participatory traditions. Participatory evaluation is much more embedded in developing countries than it is in the United States. We're getting the benefit in some of our own inner-city communities of what's being learned about participatory evaluation in international contexts.

Nor is the diversity just international. I've been working with some Native Americans on using a Navajo cosmic framework for conceptualizing and facilitating evaluation. Their emphasis is on interconnected circular and cyclical patterns rather than linear, deterministic approaches. That turns out to have lots of implications for thinking about and understanding causality and outcomes.

Jean: People sometimes speak of the hegemony of North American evaluation processes and thinking. Do you think that's true?

Michael: That perception stems in part, I suspect, from the fact that United States authors dominated much of the early literature. We have more people who write textbooks and who publish in the journals, and we have more training available, and there are more Americans doing the training.

That's shifting. It's been shifting over the last 10 years, but it's not yet as visible as it will be. Now there are international journals and international publishing outlets and international societies, and they're developing indigenous evaluators.

The United States also has a consulting industry—of which I am a part—unlike any anywhere else in the world. So we have a lot of people available to do international work and have figured out how to get those contracts. Americans dominate the contract work internationally because we have people who are set up to do it and research companies, the big ones—the RANDs, the Abts, the Urban Institutes, and on and on—who do that kind of work. So the business side of evaluation is dominated by Americans who are very entrepreneurial about evaluation. And it's a much, much bigger industry than what is visible on the academic side of evaluation.

In the United States, hegemony flows from federal government administrators, especially OMB (Office of Management and Budget) political appointees, who mandate systems like GPRA (Government Performance Results Act) and PART (Program Assessment Rating Tool) with little or no understanding of evaluation, operating on simple and simple-minded notions of performance measurement while thinking they're doing evaluation—and telling the public they're ensuring accountability. This involves millions and millions of dollars spent in complying with federal reporting mandates that are largely meaningless and useless.

And as long as you've raised the topic and given me a chance to vent, while American hegemony may be a concern within the international profession, in terms of the sheer volume of evaluation done in the world, the biggest problem, it seems to me, is that most evaluations are still done by people who have no evaluation training, don't know there is a field of evaluation, don't know there are any standards, and don't know what we've learned about making evaluations useful. Consultants and academics get contracts to do site visits for a huge number of development projects around the world, often with little or no real input from local people and intended beneficiaries and no sense of intended use. Internationally, that is a major part of what people experience as evaluation and associate with evaluation because that's what happens to them.

I mention these concerns because, as a profession, evaluators, talking among ourselves, can become convinced that we're making a big difference. I suspect there is more evaluation done now than at any time in history—and more of it is lousy and useless than at any time in history. But not everyone would share my criteria for what constitutes useful and meaningful evaluation. One thing I keep pushing for, as does AEA and other professional evaluation associations and leaders, is to be sure that people with professional evaluation backgrounds and knowledge participate in international evaluation teams and in the design of federal evaluation processes. We have professional evaluators working at the ground level trying to make these systems meaningful and useful, but we've had less influence in the overall design of such mammoth, resource-sucking, top-down, mandated systems.

On the other side of the ledger, there is strong and growing philanthropic support for evaluation and many examples of excellent practice both domestically and internationally. It's just that when I look at where most of the money is spent on evaluation, the influence of the profession can scarcely be underestimated.

Jean: We know that you love the Grand Canyon and spend a lot of time there. Probably some of the time that you're there, you do think about evaluation. If you would, please complete the metaphor for us: How is evaluation like the Grand Canyon?

Michael: I love that question and actually wrote about it in the last edition of *Utilization-Focused Evaluation* as an introduction to process use. It also relates to this notion of evaluation under conditions of uncertainty in dynamic and emergent conditions. One of the things that distinguishes people in the hiking world, what might be called a "paradigm debate" among people who spend time in the wilderness, is the folks who are process-oriented versus those who are goal-oriented. There are people who set out to get from point A to point B, to log a certain number of miles, to have hiked every mile of the Canyon, to have covered it from one end to the other, to do it with certain speed, to do the Mount Everest ascent—there is a conquering, goal-oriented side to wilderness experiences. And then there are those folks who go out to experience the wilderness and see what it gives them. No specific goals. Often, only a vague itinerary, one with several

options along the way. The process is open. You take away from it what you take away from it. That's how I experience the Canyon. I have done goal-oriented hiking in the Canyon, especially when I've done coming-of-age initiation ceremonies with my kids, using the Grand Canyon as our place of connection to nature. But even those were open in the sense that we set out to spend time together in the rugged, demanding, and beautiful environment of the Canyon, knowing that important things would happen if we opened ourselves to that experience. And important things did happen, which led to the book I wrote called *Grand Canyon Celebration* (Patton, 1999), which recounts those experiences.

When I talk about bringing complexity science to evaluation and social innovation, I'm talking about using evaluative thinking to support people who are on a journey without a clear, pre-determined destination. It's an important journey driven by vision and values, but they don't have performance objectives or measurable outcomes. Indeed, performance objectives would get in the way, would actually undermine openness and emergence. The most powerful experiences I've had in the Canyon have been what complexity scientists would call emergent experiences that you couldn't set out to try to have because you don't even know they exist. The people who are changing the world in major ways are often those kinds of people. They aren't the performance measurement people or the performance targets people. That orientation works well for evaluating immunization campaigns, but we don't know how to immunize people against poverty and social injustice. The people who are operating out of vision, social innovators who are learning to pay attention to their environment and what's going on around them and acting responsively, they want the rigor of evaluative thinking that developmental evaluation offers but without the baggage of forced, imposed, and premature clear, specific, and measurable objectives.

Figuring out how to make evaluation useful in complex, dynamic environments is where I'm putting my energy (see Westley, Zimmerman, & Patton, 2006). Part of my legitimacy with those kinds of people comes from sharing Grand Canyon stories. These are highly creative people who are often suspicious of evaluators because they experience evaluators as narrow, negative, uncreative, and constipated in their thinking. When they find out how I've experienced the Grand Canyon, what I've gotten from the Canyon, how I hike the Canyon, and how I evaluate what I take away from my Canyon experiences, they say, "That's the kind of evaluation I want."

Jean: Thanks, Michael, for taking the time to speak with us.

References

- Patton, M. Q. (1981). *Creative evaluation*. Beverly Hills, CA: Sage.
- Patton, M. Q. (1982). *Practical evaluation*. Newbury Park, CA: Sage.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (1999). *Grand Canyon celebration: A father-son journey of discovery*. Amherst, NY: Prometheus Books.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Preskill, H., & Caracelli, V. (1997). Current and developing conceptions of use: Evaluation Use TIG survey results. *American Journal of Evaluation*, 18(3), 209-226.
- Westley, F., Zimmerman, B., & Patton, M. Q. (2006). *Getting to maybe: How the world is changed*. Toronto, Canada: Random House.