

The Challenges of Making Evaluation Useful

Michael Quinn Patton, Ph.D *

Excerpt

The evaluation profession has been studying ways of increasing the use of evaluations. Here are some of the things we've learned are important for evaluations to be useful: being clear about intended uses by primary intended users; creating and nurturing a results-oriented, reality-testing culture that supports evaluation use; collaboration in deciding what outcomes to commit to and hold yourselves accountable for; making measurement of outcomes and program processes thoughtful, meaningful, timely, and credible – and integrated into the

program; using the results in a visible and transparent way, and modeling for others serious use of results.

Michael Quinn Patton, Ph.D
Faculty member of the Union
Institute Graduate School,
Minneapolis, MN, USA
Founder and Director of an
organizational development
consulting business "Utilization
- Focused Evaluation"
Former President of the
American Evaluation
Association

Trecho Extraído do Artigo

A avaliação como profissão vem estudando meios de ampliar a utilização das avaliações. Aqui são apresentadas algumas lições aprendidas como importantes, para avaliações serem úteis: clareza sobre quais são as utilizações pretendidas pelos principais usuários pretendidos; criação e fortalecimento de uma cultura de orientação de resulta-

* Michael Quinn Patton é doutor em sociologia pela Universidade de Wisconsin, Estados Unidos, consultor de avaliação há 30 anos, dirige atualmente a Utilization-Focused Evaluation, empresa de consultoria em desenvolvimento organizacional. Avaliou uma série de programas em áreas tão diversas quanto saúde, serviços humanos, educação, cooperação, meio ambiente, agricultura, emprego, treinamento, desenvolvimento de liderança, alfabetização, infância e educação para pais, redução da pobreza, desenvolvimento econômico e advocacia. Sua prática em consultoria inclui avaliação de programas, planejamento estratégico, resolução de conflitos, desenvolvimento de pessoal e uma variedade de projetos de desenvolvimento organizacional. É autor de vários livros de referência sobre avaliação de programas, incluindo *Utilization-Focused Evaluation: the new century text* (1977) e *Qualitative evaluation and research methods* (2001).

dos e constatação da realidade que apóie a utilização da avaliação; colaboração nas decisões sobre com quais resultados se comprometer e se responsabilizar; garantia de que a verificação dos resultados e do processo de desenvolvimento do programa seja reflexiva, significativa, oportuna, confiável e integrada ao programa; utilização dos resultados de um modo visível e transparente servindo como exemplo para outras utilizações responsáveis de resultados.

As we gather today here in Rio de Janeiro to consider the challenges of making evaluation useful, a good place to begin is with then observation that Evaluation, as a profession, has been internationalized. In 1995 evaluation professionals from 61 countries around the world came together at the First International Evaluation Conference in Vancouver, British Columbia. Evaluation associations from the United States, Canada, the United Kingdom, and Australia/New Zealand sponsored that conference. In 2005 the Second International Evaluation Conference will be held in Toronto. In the last 10 years more than 40 new national evaluation associations and networks have emerged worldwide including the European Evaluation Society, the African Evaluation Society, the International Organization for Cooperation in Evaluation, and the International development Evaluation Association. Brazil has its own Brazilian Evaluation Network and in October of this year the Latin American Evaluation Association will be launched in Peru.

These globally interconnected efforts made it possible for information to be shared about strategies for and approaches to evaluation worldwide. Such international exchanges also offered alternative evaluation criteria. Thus, the globalization of evaluation supports our working together

to increase our international understanding about factors that support program effectiveness and evaluation use.

The international nature of evaluation poses significant cross-cultural challenges in determining how to integrate and connect evaluation to local issues and concerns. One of the activities being taken on by new national evaluation associations is review of the North American Standards for evaluation. Those standards call for evaluations to be useful, practical, ethical, and accurate. Annex (see the end of the speech) the summary of these standards.

Fear of Evaluation and Resistance to Using Evaluation

Since the standards call for evaluations to be useful, we are faced with the challenge of how to conduct evaluations in such a way that they are useful. Resistance is so common that evaluators have created their own version of the Genesis story to describe this challenge. It goes like this:

In the beginning God created the heaven and the earth.

And God saw everything that he made. "Behold," God said, "it is very good." And the evening and the morning were the sixth day.

And on the seventh day God rested from all His work. His archangel came then unto Him asking, "God, how do you know that what you have created is 'very good'? What are your criteria? On what data do you base your judgment? Just exactly what results were you expecting to attain? And aren't you a little close

to the situation to make a fair and unbiased evaluation?"

God thought about these questions all that day and His rest was greatly disturbed. On the eighth day God said, "Lucifer, go to hell."

Thus was evaluation born in a blaze of glory [. . .].

This is what we would call a "summative evaluation" in the language of evaluation, that is, an evaluation aimed at an overall judgment of merit or worth. The traditional form of evaluation has involved having independent, external evaluators judge the effectiveness of a program, what we have come to call "summative evaluation" Summative evaluations are aimed at making decisions about which projects to continue, which to change, and which to terminate.

Formative Evaluation

A second type of evaluation is also very important, what we call "formative evaluation" or "evaluation for improvement." — an approach to evaluation that emphasizes learning, improvement, and identification of strengths and weaknesses, especially from the perspective of project beneficiaries. In such an approach to evaluation the staff, beneficiaries, and evaluator work together collaboratively to learn how to be more effective and to change things that are not working effectively. For a long time I've been in search of a creation story that had a formative element to it, and it wasn't until I came here to New Zealand that I found that story. The Maori story of creation tells of Rangi the Sky Father and Papa the Earth Mother, existing intertwined and

having their children between them, close together that their fierce embrace shut out the light and so their children lived in darkness.

The god-children became disgruntled with the darkness and conspired to separate their parents, to separate the Sky Father from the Earth Mother so as to allow the light and the wind to come in, and they did so.

Tane who was the god of the forest, of the animals, of the insects and birds, and all living things, was the most powerful of the god-children. It was finally his strength, added to that of his brothers that managed to separate the Sky Father from the Earth Mother, and upon separating them, they saw them in the light for the first time. They saw Rangi the Sky Father with tears in his eyes, that became the rain and Papa the Earth Mother in her nakedness.

And so Tane decided to clothe and decorate his mother and began planting trees in the earth mother to adorn her, but because he was young and inexperienced and still learning, he planted the trees wrong side round. He put the leaves in the earth, instead of the roots. When he had done this he stood back and looked at his handiwork and saw that no birds came and that no animals came and that it was not very beautiful. So he reflected on this, went back, took the trees out, turned them around and planted them roots first, with the leaves in the air. And immediately the birds appeared and the animals and people came out to live beneath the trees.

And so here we have a creation story of formative evaluation, of trying something out, seeing whether or not it works, and when finding that it did not work, changing the practice to make it work. The first formative evaluation creation story I've encountered.

Formative evaluation is usually done in collaboration with program staff while summative evaluation is typically done independently and externally. The staff within a program may undertake formative evaluation for internal improvement, but for accountability purposes externally, a summative evaluation would be conducted independently. Formative evaluation is also done to get ready for summative evaluation.

One of the lessons that the field of evaluation has learned is that new and innovative projects need time to develop before they are ready for summative evaluation. Damage can be done to such projects if external evaluation is done too soon. So, the typical strategic approach is to engage in formative evaluation to learn and improve, and then at a certain point when the project is ready for a more rigorous, independent process, to undertake a summative evaluation.

Another distinction important to evaluation is to separate personnel evaluation (evaluation of staff performance) from program or project evaluation (evaluation of project effectiveness). These two are quite different approaches to evaluation. One of the challenges to useful evaluation is helping intended users understand the different kinds of evaluations and the different ways to use them appropriately.

Reality-testing

All serious evaluation requires an attitude that we have come to call "a willingness to

engage in reality-testing." Reality-testing means the willingness to look at what is really going on and what is actually happening, not just what we hope is happening.

We know from the fields of psychology and other social sciences that as human beings have a tendency to believe what we want to believe. We have a tendency to see things in positive ways because we have high hopes that what we are doing is good. But we also have the capacity to fool ourselves, that is, to convince ourselves that good things are happening when, in fact, they are not. Thus, the mental mind-set for evaluation is the willingness to face reality. The mechanisms, procedures, and methods of evaluation are aimed at helping us test reality. This kind of reality-testing mentality must be established within programs so that staff, directors, and funders of programs are willing to admit and learn from mistakes. We have a saying in United States that we learn the most from our mistakes, not from our successes. In order to learn from our mistakes we must be able to recognize and admit when we have made mistakes. Evaluation contributes by helping us identify what is working and what is not working, and then to change things that are not working.

Evaluations help overcome reality distortion. Rio de Janeiro takes its name from a reality distortion. The original Portuguese explorers thought they had sailed into a river when they entered what is now Rio de Janeiro. Instead, it turned out to be a harbor, but the name stuck – an inaccurate name.

One of the important contributions of evaluation is to help program staff and planners move beyond the reality distortions of personal "intuition." Many times when I

am interviewing staff about how they know that intended beneficiaries are really being helped, they reply: "Because I feel people are being helped." That is reliance on "intuition," but the challenge of evaluation is to go beyond "intuition" by establishing concrete evidence that people are being helped.

Let me share an example. I had the opportunity in Japan to visit the Morioka Citizen's Welfare Bank which was engaged in a project to collect used clothing and used household goods from Japanese citizens to send to needy people in the Philippines. They had collected a large amount of discarded goods for donation to the Philippines. However, when this initiative was implemented, they found that many of the poor people in the Philippines were insulted to by being given discarded clothing and goods from Japan. For them, it was a loss of face to receive something that Japanese people no longer wanted. This kind of charity was experienced as demeaning. Thus, this very kind idea, this very good idea for helping fellow citizens in another country, simply did not work in practice. It took great insight and courage, I believe, to admit that this project was not working as originally designed and to change it. The project was redesigned as a technical assistance and development effort to help the people in the Philippines design their own goodwill program to collect and distribute used clothing and goods within the Philippines. Thus, this project changed from one of distributing used Japanese goods to one of building capacity within the Philippines for the collection and distribution of used goods. This is an example of formative evaluation, of the willingness of the people within an NPO program to look honestly and courageously at what was working and what was not working, and to change what was not working.

The final approach to evaluation I would like to mention is called "Knowledge-Generating Evaluation" or "Evaluation To Learn Lessons." Both summative and formative evaluation are aimed at assessing the effectiveness of specific, individual projects. In contrast, knowledge-generating evaluation looks for patterns of effectiveness across many different projects in order to learn general lessons about what works and doesn't work. For example, if we do evaluations of many different recycling projects, we can bring together the findings from those different projects to learn lessons about what is most effective in undertaking recycling initiatives. Several philanthropic foundations in United States have made this form of evaluation a priority for future evaluative studies.

Major Challenges of Making Evaluation Useful

The evaluation profession has been studying ways of increasing the use of evaluations. Here are some of the things we've learned are important for evaluations to be useful.

- Being clear about intended uses by primary intended users.
- Creating and nurturing a results-oriented, reality-testing culture that supports evaluation use.
- Collaboration in deciding what outcomes to commit to and hold yourselves accountable for.
- Making measurement of outcomes and program processes thoughtful, meaningful, timely, and credible – and integrated into the program.
- Using the results in a visible and transparent way, and modeling for others serious use of results.

Utilization-Focused Evaluation

One major approach to evaluation is Utilization-Focused Evaluation (PATTON, 1997), which builds on the observations above. It is designed to address these challenges. Utilization-Focused Evaluation (UFE) begins with the premise that evaluations should be judged by their utility and actual use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything that is done, from beginning to end, will affect use. Use concerns how real people in the real world apply evaluation findings and experience the evaluation process. Therefore, the focus in utilization-focused evaluation is on intended use by intended users. Since no evaluation can be value-free, utilization-focused evaluation answers the question of whose values will frame the evaluation by working with clearly identified, primary intended users who have responsibility to apply evaluation findings and implement recommendations.

Utilization-focused evaluation is highly personal and situational. The evaluation facilitator develops a working relationship with intended users to help them determine what kind of evaluation they need. This requires negotiation in which the evaluator offers a menu of possibilities within the framework of established evaluation standards and principles.

Utilization-focused evaluation does not advocate any particular evaluation content, model, method, theory, or even use. Rather, it is a process for helping primary intended users select the most appropriate content, model, methods, theory, and uses for their particular situation.

Situational responsiveness guides the interactive process between evaluator and primary intended users. A utilization-focused evaluation can include any evaluative purpose (formative, summative, developmental), any kind of data (quantitative, qualitative, mixed), any kind of design (e.g., naturalistic, experimental), and any kind of focus (processes, outcomes, impacts, costs, and cost-benefit, among many possibilities). Utilization-focused evaluation is a process for making decisions about these issues in collaboration with an identified group of primary users focusing on their intended uses of evaluation.

A psychology of use undergirds and informs utilization-focused evaluation: intended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings; they are more likely to understand and feel ownership if they've been actively involved; by actively involving primary intended users, the evaluator is training users in use, preparing the groundwork for use, and reinforcing the intended utility of the evaluation every step along the way.

Situational Responsiveness

Situational factors that affect evaluation design and use include program variables (e.g., size, complexity, history) evaluation purposes, (formative, summative), evaluator experience and credibility, intended users, politics, and resource constraints. An evaluator demonstrates situational responsiveness when strategizing how various factors affect evaluation design and use. The implication is that no one best evaluation design exists, that is, no

standardized cookie cutter approach can be applied regardless of circumstances and context. The standards and principles of evaluation provide direction, but every evaluation is unique.

Situational responsiveness involves negotiating and designing the evaluation to fit the specific intended uses of the evaluation by particular intended users.

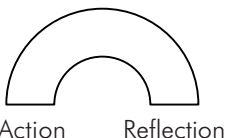
Evaluative-thinking as a Methodology and Tool

We tend to think of methodology as referring to techniques of data collection and analysis. Mixed methods evaluators often use the metaphor of a “tool kit” to remind evaluators to pick the right tool for the job. Experienced tool users remind us that having only one tool is both dangerous and limiting: famously, if all you have is off hammer, everything looks like a nail. But the focus on evaluation as a source of applied social science tools or methods remains limited.

Webster’s New World Dictionary (1995) defines methodology as “the science of method, or orderly arrangement; specifically, the branch of logic concerned with the application of the principles of reasoning to scientific and philosophical inquiry.”

This definition directs our attention beyond data collection and analysis techniques to evaluative thinking.

Evaluative Thinking



Evaluative thinking is one way to think about the connection between action and reflection. In order to reflect on action, program staff and social innovators must also know how to integrate and use feedback, research, and experience, that is, to weigh evidence, consider inevitable contradictions and inconsistencies, articulate values, interpret findings, and examine assumptions, to note but a few of the things meant by “thinking evaluatively.”

The capacity to think astutely is often undervalued in the world of action and taken for granted in the world of scholars. In contrast, Philosopher Hannah Arendt (1968) identified the capacity to think as the foundation of a healthy and resilient democracy. Having experienced totalitarianism, then having fled it, she devoted much of her life to studying it and its opposite, democracy. She believed that thinking thoughtfully in public deliberations and acting democratically were intertwined. Totalitarianism is built on and sustained by deceit and thought control. In order to resist efforts by the powerful to deceive and control thinking, Arendt believed that people needed to practice thinking. Toward that end she developed “eight exercises in political thought.” She wrote that “experience in thinking... can be won, like all experience in doing something, only through practice, through exercises.”

From this point of view, might we consider every evaluation an opportunity for those involved to practice thinking? In this regard we might aspire to have evaluation of social programs do what Arendt (1968, p. 14-15) hoped her exercises in political thought would do, namely, “to gain experience in how to think,” in this case, how to think about and evaluate the complex dynamics of social innovation in order to learn and

increase impact. Her exercises “do not contain prescriptions on what to think or which truths to hold,” but rather on the act and process of thinking. For example, she thought it important to help people think conceptually, to “discover the real origins of original concepts in order to distill from them anew their original spirit which has so sadly evaporated from the very keywords of political language— such as freedom and justice, authority and reason, responsibility and virtue, power and glory—leaving behind empty shells [...]”. Might we add to her conceptual agenda for examination and public dialogue such terms as outcomes and impacts, and accountability and learning, among many possibilities.

Helping people learn to think evaluatively can be a more enduring impact from an evaluation than use of specific findings generated in that same evaluation. Findings have a very short ‘half life’ - to use a physical science metaphor. They deteriorate very quickly as the world changes rapidly. Specific findings typically have a small window of relevance. In contrast, learning to think and act evaluatively can have an ongoing impact. The experience of being involved in an evaluation, then, for those actually involved, can have a lasting impact on how they think, on their openness to reality-testing, on how they view the things they do, and on their capacity to engage in innovative processes.

Discovering Process Use

Trying to figure out what’s really going on is, of course, a core function of evaluation. Part of such reality testing includes sorting out what our profession has become and is becoming, what our core disciplines are, and what issues deserve our attention. I have

spent a good part of my evaluation career reflecting on such concerns, particularly from the point of view of use, for example, how to work with intended users to achieve intended uses, and how to distinguish the general community of stakeholders from primary users so as to work with them. In all of that work, and indeed through the first two editions of *Utilization-Focused Evaluation* (a period spanning 20 years), I’ve been engaging in evaluations with a focus on enhancing utility, both the amount and quality of use. But, when I went to prepare the third edition of the book, and was trying to sort out what had happened in the field in the ten years since the last edition, it came to me that I had missed something.

I was struck by something that my own myopia had not allowed me to see before. When I have followed up my own evaluations over the years, I have enquired from intended users about actual use. What I would typically hear was something like: “Yes, the findings were helpful in this way and that, and here’s what we did with them.” If there had been recommendations, I would ask what subsequent actions, if any, followed. But, beyond the focus on findings and recommendations, what they almost inevitably added was something to the effect that “it wasn’t really the findings that were so important in the end, it was going through the process.” And I would reply: “That’s nice. I’m glad you appreciated the process, but what did you really do with the findings?” In reflecting on these interactions, I came to realise that the entire field has narrowly defined use as use of findings. We have thus not had ways to conceptualise or talk about what happens to people and organisations as a result of being involved in an evaluation process: what I have come to call ‘process use’.

The Impacts of Experiencing the Culture of Evaluation

I have defined process use as relating to and being indicated by individual changes in thinking and behaving that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process. Changes in program or organizational procedures and culture may also be manifestations of process impacts.

One way of thinking about process use is to recognize that evaluation constitutes a culture, of sorts. We, as evaluators, have our own values, our own ways of thinking, our own language, our own hierarchy, and our own reward system. When we engage other people in the evaluation process, we are providing them with a cross-cultural experience. They often experience evaluators as imperialistic, that is, as imposing the evaluation culture on top of their own values and culture - or they may find the cross-cultural experience stimulating and friendly. But in either case, and all the spaces in-between, it is a cross-cultural interaction.

Those new to the evaluation culture may need help and facilitation in coming to view the experience as valuable. One of the ways I sometimes attempt to engage people in the value of evaluation is to suggest that they may reap personal and professional benefits from learning how to operate in an evaluation culture. Many funders are immersed in that culture. Knowing how to speak the language of evaluation and conceptualize programs logically are not inherent 'goods,' but can be instrumentally good in helping people get the things they

want, not least of all, to attract resources for their programs. They may also develop skills in reality-testing that have application in other areas of professional and even personal life.

This culture of evaluation, that we as evaluators take for granted in our own way of thinking, is quite alien to many of the folks with whom we work at program levels. Examples of the values of evaluation include: clarity, specificity, and focusing; being systematic and making assumptions explicit; operationalizing program concepts, ideas and goals; distinguishing inputs and processes from outcomes; valuing empirical evidence; and separating statements of fact from interpretations and judgments. These values constitute ways of thinking that are not natural to people and that are quite alien to many. When we take people through a process of evaluation - at least in any kind of stakeholder involvement or participatory process, they are in fact learning things about evaluation culture and often learning how to think in these ways. Recognizing this leads to the possibility of conceptualizing some different kinds of process uses, and that's what I want to turn to now

Learning to Think Evaluatively

"Process use," as I've said, refers to using evaluation logic and processes to help people in programs and organizations learn to think evaluatively. This is distinct from using the substantive findings in an evaluation report. It's equivalent to the difference between learning how to learn versus learning substantive knowledge about something. Learning how to think evaluatively is learning how to learn. I think that facilitating learning about evaluation

opens up new possibilities for positioning the field of evaluation professionally. It is a kind of process impact that organizations are coming to value because the capacity to engage in this kind of thinking has more enduring value than a delimited set of findings, especially for organizations interested in becoming what has come to be called popularly "learning organizations." Findings have a very short 'half life' - to use a physical science metaphor. They deteriorate very quickly as the world changes rapidly. Specific findings typically have a small window of relevance. In contrast, learning to think and act evaluatively can have an ongoing impact. The experience of being involved in an evaluation, then, for those stakeholders actually involved, can have a lasting impact on how they think, on their openness to reality-testing, and on how they view the things they do. This is one kind of process use - learning how to think evaluatively.

Reality-Testing and Evaluation

My host and Guardian Angel while I've been in Brazil has been Thereza Penna Firme. She loves stories and metaphors for evaluation, so as a way of acknowledging her contributions and thanking her for her hospitality, I want to close with an evaluation story I know she likes, one from literature, the story of Don Quixote, the Man of La Mancha. In the story, Don Quixote, in his old age, loses touch with ordinary reality and conceives the project of becoming a Knight Errant, riding through the world making the world a better place. As the story unfolds, it becomes a story of different realities. Don Quixote thinks he is fighting a great army, others see it as a mere herd of sheep. Don Quixote thinks he is fighting a giant, while others see it as a windmill. Don Quixote thinks he is saving

a fair damsel while others see her as a common prostitute. Near the end, tricked by his son-in-law and the village priest to look at his own reflection in a mirror, Don Quixote enters back into the ordinary reality of those around him. But the disappointment demoralizes and exhausts him. Knowing that he is dying, those around him attempt to comfort him by assuring him that at least he will die in touch with reality. He responds with one of the great soliloquies of theater.

Reality. Life is it is. I've lived many years now and I've seen life as it is: pain, misery cruelty beyond belief. I've heard the voices of God's noblest creatures moan from the filth in the streets. I've been a soldier and a slave. I've seen my comrades die in battle or fall more slowly under the lash in Africa. These were men who saw life is it is and they died despairing. No glory. No bray of last words. Only the look of despair in their eyes questioning "Why?" I do not believe they were asking why they were dying, but why they had ever lived. Who knows where madness lies? Perhaps to be too practical is madness, to seek treasure where there is only trash, to surrender one's dreams, this may be madness. But maddest of all is to see life only as it is and not also as it should be and could be.

Evaluation challenges us to do reality testing but not as an end in itself. We examine reality so as not to deceive ourselves and to better direct our efforts at effectively creating life as it should be and could be for the intended beneficiaries of programs.

Referências bibliográficas

ARENDDT, H. *Between past and future: eight exercises in political thought*. New York: Viking Press, 1968.

NEUFELDT, V. (Ed.). *Webster's new world dictionary*. New York: London: Pocket Star Books, 1995.

PATTON, M. Q. *Qualitative research and evaluation methods*. 3rd. ed. Thousand Oaks, Calif.: Sage Publications, 2001.

_____. *Utilization-focused evaluation: the new century text*. 3rd ed. Thousand Oaks, Calif.: Sage Publications, 1997.

ANNEX

STANDARDS FOR EVALUATION

UTILITY

The Utility Standards are intended to ensure that an evaluation will serve the practical information needs of intended users.

FEASIBILITY

The Feasibility Standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

PROPRIETY

The Propriety Standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

ACCURACY

The Accuracy Standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the feature that determine worth or merit of the program being evaluated.

For the full set of the detailed standards, see: <http://www.wmich.edu/evalctr/jc/>