Reprinted with permission from Elsevier Science
Evaluation Practice, Vol. 13, No. 2, 1992, pp. 125 134. ISSN: 0886-1633

Industrial Productivity: New Perspectives for a Traditionally Trained Evaluator[1]

JONATHAN A. MORELL

OBJECTIVE

The objective of this paper is to provide a personal account of how my work in evaluation has changed over the years and how, despite the changes, essential similarities remain. I began my career in evaluation in a rather traditional mode, trained at Northwestern in the days of Donald Campbell, Lee Sechrest, Bob Boruch, Tom Cook, Paul Wortman and the rest of the original Northwestern evaluation group. My focus at the time was on human services and substance abuse. Now I am evaluating the impact of computer networks that tie manufacturing floor information to the work of product designers, marketers, and a company's parts suppliers. My work was recognized as evaluation then, and it can be recognized as evaluation now.

Evolution from Human Services to Industrial Productivity

I spend most of my time working on projects aimed at improving the productivity of manufacturing in the United States. Although not exclusively involved with information technology, this focus is consuming an ever greater percentage of my time. My work is usually aimed at bringing about change in individual organizations rather than in social policy, legislation, or the economy. As an example, I do not evaluate the impact of educational reform efforts on the preparedness of the labor force. I have, however, evaluated the impact of a worker literacy program in a steel plant.

Given the degree of change I have undergone, three questions arise.

> Why do I still consider myself an evaluator?
> Why am I still interested in AEA?
> What are the implications of my interest for evaluation and for AEA?

To answer these questions, I have to explain a few things about myself and about my beliefs concerning the field of evaluation. In so doing, I hope to show that evaluation as a field and AEA as an organization can legitimately encompass more work in my areas of interest. Whether it should do so, and whether it will do so is a matter for the leadership and membership of AEA to decide.

EVALUATION IN THE SERVICE OF INDUSTRIAL PRODUCTIVITY

To set the stage, I will briefly describe some of the projects I and my colleagues are doing.

Michigan Manufacturing Technology Center Evaluation: Think of the difference between the Chrysler Corporation, on the one hand, and the "Joe Blow Metal Stamping Works," with 200 employees on the other. Aside from the obvious differences of size and money, these companies differ in other critical aspects. "Joe Blow" has older and less automated manufacturing equipment. The sophistication of his installed base of information technology is lower. The extent of his in-house expertise is less. Joe's access to capital for improvements is more limited. As a result of these factors, the Joe Blow

[1] An earlier version of this paper was presented as an invited address to the Topical Interest Group on Evaluation in Business and Industry of the American Evaluation Association, Chicago, IL, November 1, 1991.

Reprinted with permission from Elsevier Science
Evaluation Practice, Vol. 13, No. 2, 1992, pp. 125 134. Copyright (c) 1992 by JAI Press, Inc. ISSN: 0886-1633

Metal Stamping Works is far less able to modernize than is the Chrysler Corporation. And yet 66 cents of every dollar that Chrysler receives for a car goes to the Joe Blow Metal Stamping Works and other companies like it- i.e., Chrysler's suppliers.

To generalize this scenario, no matter how much an original equipment manufacturer attempts to modernize, efficiency and quality are limited by the efficiency and quality of a supplier base that is technologically backward and difficult to help. Not only is the Joe Blow company in trouble, but, in the aggregate, so is the entire manufacturing base of the United States, along with the good jobs and large income multipliers that go into every dollar of manufacturing wages.

To address the problem of modernizing the supplier base, the National Institute of Standards and Technology (NIST) has funded ITI as one of five regional modernization centers. The level of our funding is significant. Over five years ITI will receive $13M from NIST, $13M from state matching funds, and an additional several million from private industry.

With the size of the project and the criticality of the problem, there is keen interest in seeing demonstrated effectiveness. Thus some of my colleagues at ITI are implementing a classic field experiment, complete with control groups, before and after measures, and causal models of variables within firms which affect success in manufacturing. The evaluation will assess hundreds of firms, stratified across various industrial sectors, and focus on several different methods of working with small manufacturing firms to help them modernize.

Air Force ManTech Enterprise Integration Program: "Enterprise integration" refers to a condition in which accurate and timely information flows smoothly within a company, and between the company and its suppliers, customers, and partners. In the integrated enterprise, all parts of the organization work cooperatively toward the satisfaction of common goals and objectives. Systems, procedures, and technologies are coordinated to minimize waste. The value of enterprise integration lies in its potential to help companies speed the development of better products and services, capitalize on opportunities, and reduce operating costs. The concept of enterprise integration is important because effective manufacturing is becoming more and more dependent on electronic communication among all parties to a manufacturing process.

To illustrate this abstract concept, consider some examples. In order to bring products to market sooner, manufacturing companies must engage in "concurrent engineering," a process by which related design teams, manufacturing engineers, marketers, financial experts, and others work together in developing a new product. In the old days manufacturers would receive specifications they could not build, and use those specifications to produce goods they could not sell. The solution is to lengthen initial design efforts by involving a greater variety of interested parties, and thereby reducing design changes and decreasing time to market. Making this process work, however, requires common CAD systems, common definitions of technical terms, common data bases, and the capacity to pass information back and forth.

Implementing enterprise integration is not just technological challenge. Working on it is a social scientist's dream come true. Human resource issues are involved in the form of training, reward systems, and job design. Organizational behavior and culture play a big role because people have to trust each other, members of different disciplines have to reach common understandings, and teams have to work effectively. Organizational structure is important because needed changes won't take place without a good strategic planning process, along with mechanisms of cooperation among divisions. Legislation and government policy also play a role. As an example, the government is working on a

Reprinted with permission from Elsevier Science
Evaluation Practice, Vol. 13, No. 2, 1992, pp. 125 134. Copyright (c) 1992 by JAI Press, Inc. ISSN: 0886-1633

common standard for the exchange of information between it and its suppliers. The policies it adopts with regard to those standards will have a strong impact on whether uniform communication systems develop in the future.

The Department of Defense (DoD) is interested in enterprise integration because of its impact on national defense. The military wants to decrease the cost and time of acquisition, insure product quality, and maintain independence from foreign suppliers. Because manufacturing for defense is so intertwined with manufacturing in general, any work done to support DoD needs will also have large impacts on civilian manufacturing.

Enterprise integration is a fine idea, but the fact is we have relatively little information about specific company needs for enterprise integration, how this complex concept should be implemented, or what impact it will actually have. Thus we have recently completed a needs assessment for enterprise integration. The methodology involved content analysis of planning documents and telephone surveys with experts. Soon we will begin working on defining metrics for an outcome evaluation so that the impact of integration in selected sites can be measured over the next few years.

Organizational Change in an Automobile Assembly Plant: The new wave in manufacturing practice includes the use of work teams, cross-functional communication, and employee involvement as means of increasing productivity, improving quality, and lowering cost. ITI did an interesting evaluation of an effort to install such a system in an automobile assembly plant. The organizational intervention was major both in terms of dollars and hours. Several full-time process consultants were on site for a year. Every one of the 5,000 personnel in the plant who agreed to participate-from the shop floor to the plant manager-was interviewed in an effort to understand precisely what that person did, and how his or her job related to the work of others. Teams met to analyze problems. Replacement workers were paid so that shop floor workers had release time to participate in the program. Funds also went to approximately twenty interviewers and four full-time coordinators. The money came jointly from union and management, and both were keenly interested in how well the program worked.

Enter ITI with a team who developed closed- and open-ended interview protocols, talked to a random sample of employees from all shifts, content analyzed suggestions that came out of team meetings, and carefully traced the impacts of suggestions that were implemented. We also tried, and failed, to develop comparison groups by exploiting lags in program implementation in different parts of the plant. Moreover, we ran into a problem that is certainly familiar to many evaluators -- changing implementation schedules that were beyond our control.

Discerning the Standards Process: Think about purchasing audio equipment. You could buy any brand or model of amplifier, turntable, CD player, and tape player, and be 100% certain that they will all connect. Plugs for the same functions are the same size and shape, as are their attendant receptacles. Input/output parameters for electronic signals are similarly standardized. Contrast this situation with an attempt to pass data back and forth among computers. If we knew nothing about the computers, operating systems, bcal area networks, and applications involved, we would all estimate the probability of successful data transfer as close to zero. These two examples differ so starkly because in one case there are agreed upon standards, and in the other case there are not. The Industrial Technology Institute is interested in this matter because effective communication is critical to modernization.

The problem of standards is ubiquitous, and the subject of work for thousands of committees around the world, all linked in a complex web of voluntary industry associations and government activity. As you

Reprinted with permission from Elsevier Science
Evaluation Practice, Vol. 13, No. 2, 1992, pp. 125 134. Copyright (c) 1992 by JAI Press, Inc. ISSN: 0886-1633

might imagine, the details of a standard are matters of intense interest to many, as they can make the difference between commercial success or failure for products and industries.

The implications for products are obvious-how many computers can you buy today with CPM operating systems? How many bicycle helmets can you find that fail to meet the recently developed ANSI standard for safety? In both cases, the answer is none and the companies that made those products either changed or departed.

What about whole industries? Think about why we do not have high definition television or good quality AM sound on our radios. These industries do not exist largely because effective standards could not be set.

The CPM example also raises an interesting complication. While some standards are set by committee, others are set by the market. No committee legislated against CPM. That operating system was done in by the market acceptance of Microsoft's Disk Operating System.

Despite the critical importance of standards for industrial development, and the fact that standards development is a complicated social, political, and economic process, there is almost no social science research that systematically evaluates how standards processes work, and why they succeed or fail. We are out to correct that lack by doing an in-depth comparative case study of the relative success of a standard called the Manufacturing Automation Protocol, which facilitates communication among devices on the shop floor. We chose this standard for two reasons. First, it was largely developed at the Industrial Technology Institute, thus providing access to much relevant information. Second, it is yet another instance where American technology finds more success overseas than at home. The U.S. is our acceptance failure" condition, while Europe represents the success case.

ITI RESEARCH STRATEGY AS EVALUATION

These examples are pieces of an overall four-pronged strategy that ITI has developed for pursuing our mission of improving the competitiveness of American Industry in the global economy. Elements of that strategy are:

Consensus development-We seek to further a broad consensus on needed technologies for modernization, on standards to support those technologies, and on methods of technology deployment. Without such a consensus, work will be unfocused and synergy among diverse initiatives will be lost.

Market-based technology transfer - only a small percentage of our huge investment in technology ends up making a discernable contribution to meeting practical needs. One reason for this problem is that effective means are lacking to identify technology development activity that will meet real needs, and to transfer that technology from its venue of development to commercial settings.

Product evaluation and validation-A product can be developed and commercialized, and still not be trusted. It is one thing to say that a product conforms to a set of standards, it is quite something else to know that the product will perform as advertised. As an example of how serious this problem can be, consider a plant manager who is asked to bet his production schedule and his performance review on the proper functioning of a new piece of software. Without appropriate product evaluation and validation, that plant manager would have very good reason not to employ the new technology.

Sub-tier deployment: I have already described the special problems of the small manufacturers who make up the bulk of our manufacturing base. The special needs of this group must be a key element in any overall modernization effort.

Reprinted with permission from Elsevier Science
Evaluation Practice, Vol. 13, No. 2, 1992, pp. 125 134. Copyright (c) 1992 by JAI Press, Inc. ISSN: 0886-1633

Actualizing this strategy requires innovative programs which may or may not work as designed. That is why all of the examples I cited involve assessing the practical consequences of trying to bring about change in real world settings-which is close to "evaluation" as I see the term.

If you strip away the unfamiliar subject matter you will see stark similarities with evaluation in human service or education. The quantitative and qualitative methods seem familiar. The design problems and difficulties with program description and causal inference seem familiar. The need to help decision makers is familiar, as is the "political" nature of the work.

Moving on from the similarities in all the evaluation work I have done over the years, let's discuss how I came to do such work, and more important, why I still feel at home in evaluation and in the AEA.

If I trace my intellectual interests, three streams were always present: innovation adoption, organizational behavior, and the use of scientific and technical information for making practical decisions. One day I had an epiphany-I could combine all my interests by studying how managers used microcomputers. I began some shoestring research on this topic, and evolved over a period of years to the state you see me in now.

RELEVANCE OF EVALUATION AS A DISCIPLINE AND THE AEA AS A PROFESSIONAL GROUP

Now for the more difficult and more interesting question. Why am I still interested in evaluation and in the AEA? To give this explanation I have to plagiarize some material I used previously.

In that work I developed the notion that "evaluation" could be characterized as a loose but enduring coalition. It is loose because its members have relatively little in common. It is enduring because what they do have in common is important, and cannot be easily obtained in other settings. I see us as a group of part-time evaluators, with diverse professional needs and a shared set of ill-defined but dearly held "core professional tenets. "

Consider some of the evidence for diversity. Journals and books that specifically cater to evaluators carry an extremely wide range of materials which vary in terms of methodology, topic area, and disciplinary foundation. It is hard to imagine a space within these dimensions that has not been applied to one or another evaluation study.

This variation is also manifest in journals which are not exclusively dedicated to evaluation, but which publish evaluation material. I got a sense of the variation by checking two volumes of the *Evaluation Studies Review Annual.*[2] Recall that the Annual's policy was to reprint the best evaluation and evaluation-related articles that the editors could find. When going through these volumes I recorded the original sources of publications that were not dedicated to evaluation. As an example, I did not count articles from *Evaluation Review,* but I did count articles from the *Journal of Educational Psychology,* and the *Journal of Chronic Diseases.*

[2] 1 had to make do with the 1983 and 1985 volumes. Theft and/or inefficiency at the University of Michigan Library precluded a larger sample.

Reprinted with permission from Elsevier Science
Evaluation Practice, Vol. 13, No. 2, 1992, pp. 125 134. Copyright (c) 1992 by JAI Press, Inc. ISSN: 0886-1633

The resulting list was instructive. A rough grouping showed six education journals, six journals in psychology or sociology, three in methodology or statistics, eight in health, and six in social policy or organizational behavior. This diversity does not exactly indicate a highly focused discipline.

Another indicator of diversity comes from looking at the institutional homes of people listed in the 1989 edition of the AEA Directory. In so doing, the following numbers emerged:

college or university-40%
state or local government-12%
non-profit organization- 11%
private business-10%
federal governments%
school system-5%

Although the representation is weighted in favor of universities, it is also obvious that evaluators have found homes across a wide spectrum of institutions in the United States.

A third indicator comes from the Directory's data on "primary discipline." Only 6% of the membership list evaluation as their primary discipline. There is a high emphasis on psychology or sociology, which makes up for 45% of the members, but it takes another 12 categories to capture the primary discipline of all of the Society's members.

A fourth indicator is the overlap between the combined authorship of *Evaluation and Program Planning and Evaluation Review,* with membership in the AEA. Presumably, both publication in those journals, and membership in the Association, represent some level of affiliation with evaluation. Yet the overlap is only 33 percent. I don't know what the overlap is for authors in psychology journals and membership in APA or APS, but I bet it's a lot higher than 33 percent.

Finally, there is minimal overlap between the authors in *Evaluation and Program Planning and Evaluation Review,* even though both publications clearly advertise themselves as being primarily related to evaluation.

Why Evaluation Endures

If evaluators are as diverse a group as I just described, and if evaluation is such a small percentage of their professional lives, why does the field endure? I think it is because evaluation offers something that is not found, or at least not easily rewarded in the main professional homes of most evaluators. Put another way, we share a set of professional values not shared by the majority of our colleagues in our parent disciplines. I believe there are three such values.

- Research should focus on assessing actions that can have practical consequences in real world settings.
- Research in all phases of its life cycle-from design to reporting of results, should aid decision and policy makers.
- The political nature of programs have legitimate (but not exclusive) claims on the conduct of evaluation.

By accepting these values we can take what we know from psychology, sociology, and so forth, and apply that knowledge to an extremely important set of problems. But in the act of that application, we leave our colleagues who are working in the scientific paradigm, and enter the realm of technology. Good scientists and good technologists work in very different ways.

Reprinted with permission from Elsevier Science
Evaluation Practice, Vol. 13, No. 2, 1992, pp. 125 134. Copyright (c) 1992 by JAI Press, Inc. ISSN: 0886-1633

Those differences extend to perspectives on causation, required levels of accuracy, and types of acceptable evidence. Consider these differences, and think about their implication for how our work is done.

> *Perspectives on causation:* The technologist will be interested in those causes of a phenomenon which are most immediate, most powerful, or most manipulable within the constraints of a specific real world context. The scientist will be most interested in studying causal factors which will clear up a conceptual difficulty, determine the truth of a speculation, or advance the development of theory. There is no reason to assume that the scientific and technical choice criteria will lead to the same priorities for research, and many reasons to believe they will not.
>
> *Levels of accuracy:* Small differences from predicted results can have large implications for theory development. Thus people engaged in scientific research must try to make their measurements as accurate as possible within the context of the problem being studied. Technologists can specify acceptable levels of accuracy because differences small enough to have no practical value need not be considered. And the definition of "practical value" is social, political, and economic-thus to be decided not by evaluators, but by relevant stakeholders.
>
> *Types of evidence:* The reason for technological investigation is often to help decision makers commit scarce resources to address significant problems. Whether it is a question of building the longest suspension bridge yet known, or privatizing the schools, or instituting a new mental health program, decision makers need advice about feasibility. Often they do not want to know what solution is best, but to know what will generally work, and which bad alternatives to eliminate. Each time a solution is shown to work at an acceptable level, the faith of decision makers in committing to a course of action is increased. Thus, in contrast to scientists, technologists must heavily emphasize the objective of *confirmation* in the work they do.

These differences are also reflected in historical analyses of the relationship between science and technology. Studies reveal a rather complicated picture of how technology and science relate. A very great deal of technology requires no scientific base. When a scientific base is required, the relation between science and technology is often characterized by time lags of varying durations, and complex knowledge transfer dynamics among disciplines and groups of people.

Two of my favorite quotes capture the totality of the differences I outlined above.

> Technologists apply science when they can, and in fact achieve their most elegant solutions when an adequate theoretical base exists for their work, but normally are not halted by the lack of a theoretical base. They fill in gaps by drawing on experience, intuition, judgment, and experimentally obtained information.
>
> The aim of technology is to be effective rather than true, and this makes it very different from science.

I know I have overdrawn the differences between science and technology, but I think the caricature rings true, and is keenly felt by evaluators who have been trained in the scientific perspectives of their disciplines.

If you accept even an approximation of the view I have drawn above, the appeal of evaluation as a discipline, and of AEA, becomes clear-they are homes for people who want to devote some portion of their professional activities to work that is not recognized or rewarded in their home disciplines.

Reprinted with permission from Elsevier Science
Evaluation Practice, Vol. 13, No. 2, 1992, pp. 125 134. Copyright (c) 1992 by JAI Press, Inc. ISSN: 0886-1633

CONCLUSION

The view of evaluation I outlined above served as an excellent guide for me when I did work in human services and substance abuse. My work setting has changed. My substantive area of expertise has changed. My funders have gone from NIMH and NIDA to the Michigan Strategic Fund, the Department of Defense, and the General Motors Corporation. My collaborators are no longer psychologists, teachers, or counselors. Rather, they are computer scientists and manufacturing engineers. And still, the same principles I outlined above serve me well. I combine an eclectic collection of methods. I care about what will work and not what can be scientifically proved. I worry about balancing the requirements of good methodology with the needs of decision makers. Thus I still seek the company of people who understand these issues, who can teach me better ways to apply them, and who can comfort me in a knowing way when I can't strike the proper balance of competing requirements for good evaluation. That is why I stay involved with the field of evaluation and with the AEA.

REFERENCES

Jarvie, I.C. (1972). In: C. Mitcham and R. Mackey (Eds.), *Technology and the Structure of Knowledge*, pp. 54-61. New York: Free Press.

Morell, J.A. (1979). *Program Evaluation in Social Research*. Elmsford, NY: Pergamon Press.

Morell, J.A. (1990). *Evaluation-Status of a Loose Coalition*. Keynote address presented to the Eastern Evaluation Research Society, Villanova, PA, May 21, 1990 (Edited version: *Evaluation Practice*, /1(3), 213-219).

Weisner, J. (1970). The need for social engineering. in: F. Korten, S. Cook and J. Lacey (Eds.), *Psychology and the Problems of Society*. Washington, DC: American Psychological Association.