

The Threat Of Evaluation In Research And Development

*William R. MacDonald, Krawetz and MacDonald
Edmonton, Alberta*

RESUME

L'évaluation des programmes, telle qu'exercée par la communauté d'évaluation, est rarement pratiquée sur les programmes de recherche et de développement (R&D). Les scientifiques et les directeurs de recherche se sentent menacés par une évaluation en bonne et due forme, et qu'ils n'entreprennent donc, ni ne prêtent appui aux évaluations. Cette menace est due à trois facteurs principaux, inhérents à la communauté scientifique ainsi qu'aux organisations en général: la mythologie de la science, l'ignorance, de même que la peur de la découverte. Le témoignage que nous apportons ici, suggère plusieurs issues à chacun des facteurs: le mythe de l'impossibilité d'évaluer les programmes de recherche, le fait que les directeurs et les scientifiques ne connaissent pas les méthodes d'évaluation, ainsi que la crainte qu'ils ressentent vis à vis des changements de programmes de même qu'envers les experts venus de l'extérieur. La situation actuelle ne saurait évoluer sans un changement fondamental au sein de la communauté R&D.

ABSTRACT

Program evaluation, as practised by the evaluation community, is seldom performed on research and development (R&D) programs. Scientists and research managers are threatened by formal program evaluation, and therefore do not initiate or support evaluations. This threat is due to three major factors inherent in the scientific community or in organizations in general: the mythology of science, ignorance, and the fear of discovery. The evidence given here suggests that each factor consists of several issues, some being the myth that research programs cannot be evaluated, that managers and scientists do not know about evaluation methods, and that they are afraid of program changes and outside evaluators. The current situation is not expected to change without a fundamental shift within the R&D community itself.

Introduction

Why do we seldom determine how successful, efficient or effective our research programs are? Why is this information not available? Unfortunately, the mystique about science and research has contributed to the resistance of scientists and research managers to program evaluation.

The analysis and premises presented here are based on a brief review of the evaluation and the science management literature, interviews with research managers in Alberta, and personal observations of the scientific community. The basic premise is that scientists and research managers are threatened by program evaluation. Three general factors contribute to this threat: the mythology of science, ignorance, and the fear of discovery. Thus scientists and their managers have not developed the necessary evaluation

methodologies, have rejected evaluation as a legitimate activity for their programs, and have not examined other research programs—when information on them can be obtained—to learn from past experience.

A recent paper provides considerable support for the ideas presented here. Blais (1986) has developed a conceptual framework to explain the phenomenon of resistance to program evaluation. He suggests that people resist evaluation essentially to protect themselves from the threat they perceive to their interests or their physical or psychological integrity. Blais divides the dynamics of resistance into three categories: behaviors people exhibit (manifestations), sources of resistance, and conditions to reduce resistance. The three factors suggested in this paper, the mythology of science, ignorance, and fear of discovery, mainly relate to the sources of resistance category that Blais proposes.

While much of the evidence presented here relates to the natural and physical science R&D community, a parallel exists with the social science community. Social scientists tend to be more oriented to examining issues relevant to understanding the practice of their affairs. Yet, Lindblom and Cohen (1979:6) "... suggest that social science and social research are only weakly understood by their own practitioners. . . (these) practitioners have not applied the tools of their trade to the trade itself." This suggests that the natural scientists will understand their trade even less; such a lack of awareness can only contribute to resistance or disinterest in evaluation.

Evaluation in R&D

Despite the above statements, evaluation is an intrinsic part of science. But program evaluation is not. A discussion paper by the Comptroller General of Canada (1986:26-32) identifies six approaches to determine the quality of research: peer review, and counting the number of publications, citations, awards, technical reports, or patents. Project evaluation models abound: "Literally hundreds of models have been developed for R&D project evaluation. . . (with) notable exceptions, most of them have never been adopted for routine use" (Souder, 1978:29). Few proposal evaluators use analytical techniques, even though they demand technical and analytical rigor in the projects they evaluate. One might observe that most decisions on funding research projects are made on the basis of negotiated discussion, gut feeling, a knowledge of the science involved, and a belief that the output of the project will be useful.

Other forms of evaluation are used in organizations. Most organizations have some form of performance appraisal of their employees, however, this tends to be directed to the employee's actions, and infrequently considers the program context. Most organizations go through an annual budgeting cycle and require some form of justification for the programs their R&D groups undertake. Again, the level of analysis in the annual budget cycle may not be suitable for what we would call program evaluation—frequently all that is required is a listing of projects that the scientists have decided should be done. Industry has another form of evaluation—the bottom line. The market acceptance of industrial products is a good measure of the suc-

cess of research, although it tells very little about the management process in getting the product to market.

Another form can be called 'pseudo-evaluation', in reality a form of criticism practiced by many scientists or science managers. Frequently it is a negative assessment of another organization, or another group in the individual's own organization. These assessments are usually based on casual external observation, and most often imply that the manager's own house is in order. Krawetz and MacDonald (1985) identified several instances of pseudo-evaluation expressed by research managers during interviews:

- "The government is funding projects which don't make sense."
- "A lot of what government departments call research is questionable."
- "Academics are out of it!"
- "(A specific program) was a mess."
- "(A specific organization) is an example of industry/government co-operation."
- "(A specific funding agency) was great to work with."
- "(A specific organization) would never have existed if (another organization) had a good reputation."

Program Evaluation in Canada

Program evaluation of R&D does occur in Canada, but evidence of it is limited. In the federal government, program evaluation is an integral component of the planning and management systems of each department. The science based departments have been developing procedures and to date have conducted several evaluations. At the 1985 CES Conference in Toronto, Mr. Michael Raynor, the Comptroller General, publically stated that program evaluation reports are public documents. Obtaining them is a different matter, and little information makes its way into the published literature. In Alberta, program evaluation is not part of the government system, consequently very little has been done, and what might have been done is not well known.

The Danger in Evaluation

One international example illustrates the point that evaluation is threatening. Martin and Irvine (1983) published a paper in which they evaluated radio astronomy facilities in England. Even before the paper was published their work had generated considerable debate, including the threat of a lawsuit by Sir Bernard Lovell. Martin and Irvine used a method called 'converging partial indicators' to measure relative productivity of matched research teams at major facilities. This included three measures: assessing the significance of the total number of papers produced by different teams, quantifying scientific impact by the frequency of citations, and interviews with scientists. Irvine "... suggests that the criticism of their work is partly motivated by self-defense, and that public critics are often more muted in private. 'These people have to defend their interests,' he says, suggesting

that 'the scientific community seems to feel threatened by external review of this type'' (Dickson, 1983:483).

The Threat of Evaluation

Three main factors that contribute to a reluctance by scientists and research managers to evaluate their programs have been identified as the mythology of science, ignorance, and the fear of discovery. Each is a collection of issues that are part of the scientific domain; some are related to science itself, while others are inherent in programs or organizations, regardless of the field (they may exist in education, social services, etc.).

Mythology of Science

The following beliefs are prominent within the scientific community, and may contribute to the lack of progress made in evaluation.

- *It is not possible to evaluate research.* The research process is supposed to depend on creativity and the unique ability of researchers to practice their craft. Additionally, a long lag time may exist between the conduct of research and use of the results. Scientists suggest research output is unpredictable, so to evaluate on the basis of what was predicted and what eventually occurred is not feasible. They also may not believe their work can be evaluated when they deal with many clients and do not have a structured program. Other arguments could be advanced, but these few statements indicate the general approach taken to resist analysis of research activities. While these statements may have merit, to accept them as universal to the research process is tantamount to saying that the practice is chaotic and unmanageable. It also negates our ability to develop suitable evaluation methodologies when science itself strives to develop methods. There may be situations where creativity, predictability, or uncertainty are paramount, however, this does not negate the ability we have to manage our endeavours.
- *Only a scientist can evaluate another scientist's work.* This argument is advanced to restrict evaluation to those individuals who practice the same discipline as the scientists being evaluated. "The barrier to evaluation set up by the scientists and engineers in R&D can be considerable. It has been noted by management researchers that research scientists in particular, tend to resist the attempts by non-scientists—managers, clients or the general public—to evaluate scientific work" (Clarke, 1986:34). But the argument should only apply to evaluation of the specific science and research methodologies being practiced. Nevertheless, this argument ignores the reality of research activities, that each project takes place within a larger context, and it is this larger context that program evaluation focusses on. Except for the basic research projects sponsored by unrestricted funding agencies, most of the research supported by industry, government, and the granting agencies can be evaluated within a program context. The science is evaluated by peer review, whereas the program needs to be evaluated on different grounds, using management principles and evaluation techniques developed for programs.

- *Idealism in science.* Finally, scientists have a mythology about the conduct of science, created by the way science is portrayed in textbooks, how the future scientist learns about the scientific community and the research process, and by the aura of mysticism promoted by the scientific professions. Texts and journals distill out the practice of research, presenting the objectified constructs of the disciplinary findings. From this we presuppose that the conduct of science is rational, linear, and objective. The reality of science is much different. One only has to read recent books on the historiography or sociology of science to get a much different picture (see, for example, Brannigan, 1981; Latour and Woolgar, 1979; Kuhn, 1970). How many physicists know that Newton was an alchemist (Berman, 1984)?

Ignorance

The word ignorance implies a lack of knowledge. From a broad perspective scientists, including research managers, are generally ignorant about management. The published literature refers to this issue; the study of research managers in Alberta (Krawetz and MacDonald, 1985) also substantiates this statement. R&D managers are seldom exposed to management knowledge—either through the literature, or through training, whether in universities or in short courses offered by consultants or extension departments at universities. Generally, management knowledge is not part of their normal knowledge acquisition process, except for that acquired through experience. The following points illustrate this issue.

In 1974/5, Twiss (1976) surveyed 57 senior R&D managers from British and Continental organizations who were taking management courses at a British university. "The paper suggests that knowledge of modern management thinking has, to date, made only a limited penetration into R&D" (Twiss, 1976:87). Twiss focussed on their knowledge of concepts and techniques at an elementary level, the application of these concepts, and on their attitudes to managerial problems. "Rarely does the R&D manager undergo any preparation for his managerial responsibilities and unlike managers in other functions of the business is not often exposed to modern management research and thinking" (Twiss, 1976:91).

Krawetz and MacDonald (1985) found similar evidence. Their interviews covered a wide spectrum of R&D in Alberta—a cross-section of government, in the natural sciences, education, social services, health, law, etc.; industry, including consulting; and the university community. Research managers ranged from the heads of laboratories to one-two person groups in a government department. Some of their findings are summarized here.

- *Management Training.* Almost all managers had technical training in the field they were managing. Of the 67 managers interviewed, only one had an M.B.A., therefore all were from non-management fields. Only 12 managers commented on their personal experience, and most of these had some management training, but not recently. One manager said he had given up on courses—courses were not worth it. Unfortunately, few R&D related management courses are available in Canada.

- *Personal Knowledge Development.* Scientists are trained to acquire new knowledge through access to the literature. Several managers were asked to identify journals and books they read, on the assumption this would indicate their management interests.

"Most of the 22 managers who discussed their reading interests rely on disciplinary trade or technical journals for information on research management. Nearly three-quarters of these read only trade or science journals related to their scientific backgrounds. Most of the remainder read some management literature, but in most cases it is only a small part of their reading. One said he would not read management journals." (Krawetz and MacDonald, 1985:20)

- *Knowing if Research is Used.* There seems to be a general climate that once research is completed the role of the manager stops, although exceptions do exist. Managers could not provide documentation of research use, nor did any indicate they identified the users and wrote reports oriented to their needs. Knowledge of use is indicated by the following statements:

"We give reports wide distribution. . . make them publically available."

"Our research staff transfer the knowledge to field staff by having close contact."

"We have an individual who acts as a vehicle for tech transfer." (Krawetz and MacDonald, 1985:15-16)

- *Accountability.* Managers were asked about evaluation, particularly program evaluation. As with research use, the evaluation process was haphazard. Project evaluation was the most common approach, both for selecting projects and evaluating the final outcome (peer review was most often used), although many groups did not do a scientific evaluation of final reports. Program evaluation was of little interest. Only two mentioned that programs were being evaluated, and the rest did not mention program evaluation from a science or management perspective, indicating they do not tend to think in program terms. Three managers indicated they were planning to evaluate a program in the future.

The Krawetz and MacDonald survey gives some indication of science managers' knowledge of management and their interest in evaluation. The survey was not designed to probe for specific knowledge, rather it obtained a general appreciation of the managers' knowledge and their interest in knowing about, understanding, and applying management techniques. Certainly, program evaluation is not a priority, let alone an interest area.

The Fear of Discovery in Evaluation

The third major area in the threat triad of mythology, ignorance and fear of discovery is that of fear itself. Fear is difficult to identify and document. However, some evidence exists that fear may be basic to the evaluation process, and may have many roots. Because of this, it does not manifest

itself overtly, but must be deduced from the actions of individuals, or recognized from the evaluations that have been conducted. Eight sources of fear have been identified.

- *The fear of a negative evaluation.* A negative evaluation has implications for a manager's future. Managers may be fearful that an evaluation will expose past inefficiency or ineffectiveness. This can be critical if this is a first evaluation, or if a program which has existed for a long time is being evaluated. Many past events may have happened that the manager fears being exposed, and time has a way of distorting the specific events, since detailed activity records are very seldom kept. A negative evaluation may require the manager to answer for his actions.
- *The fear of change.* "For most program officials there is a certain comfortableness in continuing to follow existing practice. Evaluation has the potential of destroying this comfort, of forcing the program official to give up his customary procedures and consider new alternatives" (Agarwala-Rogers, 1977:330). All individuals have a style of management in which they will feel most comfortable; they will use this in most situations they encounter, even if it is not appropriate. Unfortunately, most managers have little understanding of the field of management, consequently they will have little awareness of the breadth of management styles, decision-making techniques, and organizational constructs available. With very few role models to learn from, they will not be aware of, or feel comfortable in exploring, new alternatives.
- *The fear of reviving old wounds.* Conflict is endemic to organizations, consequently, conflict probably occurs in most research programs. The management literature does discuss ways to manage conflict, but research managers are not aware of this information. A program evaluation may have to analyze and review the events surrounding the conflict situations, because of their impact on decision making, yet the participants will want to downplay what happened.
- *The fear of finding little impact from the program.* "The very reasonable expectation that a good (evaluation) study will show no effect (of the program) is itself a major reason why administrators fear hard-headed evaluation. . . . It is a very reasonable fear because the null effect is going to be interpreted as the administrator's failure or an error in his advocacy of the programme" (D. Campbell, in Salasin, 1973). Most managers interviewed in the Alberta survey believed their research output was used, but no one produced evidence that examined this use. There is an unfounded belief that the production of research output is an end in itself. Because managers do not have a feel for the success of their research efforts, they may have a deep-seated fear that there really has been little effect; for this to be discovered can be very threatening.
- *Fear of vagueness in the program.* Rutman (1980) has pointed out that a program may not have established suitable objectives and goals. "The manager is placed in the position of acknowledging vague goals. . . . Goals that merely serve symbolic or political purposes are being sorted out from

those that the manager actually pursues. The specification of goals can actually result in the questioning of how a program is being managed. This is especially threatening for programs where officials have considerable discretion in exercising their responsibilities" (Rutman, 1980:98). In programs in education, social services, and health services, evaluators very often find inappropriate goals, activities which do not relate to goals—but which may be more appropriate—and the non-pursuit of established goals.

- *Fear of an outside evaluator.* "Being an outsider ensures that the evaluator comes to his task with fresh ideas and a wide perspective, and that the evaluation results will be perceived as having relatively higher credibility than if the evaluator were an insider" (Agarwala-Rogers, 1977:330). The manager may fear that he will have little control over the outsider, and be unable to hide things from him or sway his thinking (this can work both ways, and the manager may indeed be able to hide things from the evaluator). The outsider, if he is a trained evaluator, will also use methodologies unfamiliar to the manager, who may question their validity in being applied to his 'unique' situation. Furthermore, if the evaluator is not considered to be a scientific peer in the area being evaluated, the manager may consider this sufficient reason to disregard the findings. Of course, the evaluation examines more than the science—evaluation should consider the whole program.
- *Fear that the program may be discontinued.* Conditions under which the program was initiated may have changed sufficiently and an evaluation may question its existence. For example, the program may actually have done enough research to meet the original needs, but the infrastructure in place—people, positions, power—has a life of its own and the manager wants to preserve it. Furthermore, the evaluation may raise questions about the need for the program in the first place: an original political reason for establishing the program may have been to deflect public opinion, and the reason may no longer be valid, yet the politicians or program creators have not acted to discontinue it.
- *The fear of evaluation, by other managers in an organization.* Fear may not lie with the research manager himself, but may reside in his superiors. Researchers are taught to be questioning, to evaluate their findings, and have experience at least with peer review. However, the same inquisitiveness may not hold for managers from different backgrounds. Thus, the manager's superiors may prevent an evaluation from proceeding, if it is seen as investigating their organization. The R&D function may also be seen as a low priority in the organization, and if the organizational culture does not include evaluation as a legitimate activity, senior managers may consider evaluation of the R&D function as a useless exercise.

Conclusions

Program evaluation is rare in R&D, if we are to base this understanding on the available literature. The analysis presented here suggests that evaluation

is threatening to scientists and research managers. Three general reasons, that of the mythology of science, ignorance of evaluation approaches, and fear of unpleasant news, are suggested as contributing to this threat. The situation is not expected to change in the near future. Only as we understand more about the resistance to evaluation, and have more people involved in program evaluation practice, particularly developing appropriate methodologies, will we see evaluation becoming more common.

References

- Agarwala-Rogers, R. Why is Evaluation Research not Utilized? *Evaluation Studies Review Annual*, 1977, 2, 327-333.
- Berman, M. *The Reenchantment of the World*. New York, NY: Bantam, 1984.
- Blais, R. Résistance et Collaboration à l'évaluation de Programme: un Cadre d'analyse et un Modèle Théorique. *The Canadian Journal of Program Evaluation*, 1986, 1(1), 47-58.
- Brannigan, A. *The Social Basis of Scientific Discovery*. Cambridge: Cambridge University Press, 1981.
- Clarke, T. *The Evaluation of R&D Programs and Personnel. A Literature Review*. Ottawa: Stargate Consultants, 1986.
- Comptroller General of Canada *Evaluation of Research and Development Programs Discussion Paper*. Program Evaluation Branch. Ottawa, March, 1986.
- Dickson, D. Study of Big Science Groups Hits Raw Nerve. *Science*, Apr. 29, 1983, 220, 482-83.
- Krawetz, N. and W. MacDonald *Research Management Issues in Alberta. A Cross-Organizational Study*. Edmonton: Krawetz and MacDonald Research Management Consultants, 1985.
- Kuhn, T.S. *The Structure of Scientific Revolutions*. Chicago, Ill: University of Chicago Press, 1970.
- Latour, B. and S. Woolgar *Laboratory Life*. Beverly Hills, CA: Sage Publications, 1979.
- Lindblom, C.E. and D.K. Cohen *Useable Knowledge*. New Haven, Conn.: Yale University Press, 1979.
- Rutman, L. *Planning Useful Evaluations. Evaluability Assessment*. Beverly Hills, CA: Sage Publications, 1980.
- Salasin, S. Experimentation Revisited: A conversation with Donald T. Campbell. *Evaluation*, 1973, 1, 7-13.
- Souder, W.E. A System for Using R&D Project Evaluation Methods. *Research Management*, September, 1978, 21, 29-37.
- Twiss, B.C. A Study of the Diffusion of Management Concepts and Techniques amongst R&D Managers. *R&D Management*, 1976, 6(2), 87-92.