

“Effectiveness: Reporting and Auditing in the Public Sector” — A Step Forward, Backward, or Sideways?

Richard W. Nutter
Faculty of Social Welfare
The University of Calgary

Effectiveness Reporting and Auditing in the Public Sector proposes that managers in the public sector should be responsible for providing effectiveness information in much the same way they are now responsible for providing financial information which is then subject to audit. These “management representations of effectiveness” would then be subjected to attest audits that would provide assurance that the information met accepted standards of accuracy, completeness, and balance or that it did not meet with those standards in important ways. If the comparison was favorable, the auditor would give a positive assurance. If negative, the auditor would presumably specify the variances from those standards.

In this context, *Effectiveness Reporting and Auditing in the Public Sector* can be viewed as an initial attempt to develop standards for the comprehensive audit of effectiveness. As such it addresses only the first phase of a three phase process to develop standards. That first phase is to gain agreement about which attributes are to be measured. The second phase would be to gain agreement about how those attributes should be measured. In this second phase the attributes must be operationally defined and substantial consensus attained regarding the appropriateness of those operational definitions. The third phase is determining the ranges of performance that are to be expected with regard to each of the attributes as measured by the agreed operational definitions.

The report defines effectiveness as composed of twelve attributes. This definition is much broader and more inclusive than the definition used by the Auditor General of Canada in auditing effectiveness. In fact, the Canadian Comprehensive Audit Foundation (CCAF) appears to have attempted to capture in its definition of effectiveness everything worth knowing about a program. There is very little included in the CCAF definition of effectiveness not also included in the definition of *comprehensive program evaluation* as offered, for example, by Peter Rossi and Howard Freeman (1985; 13, 19, 38).

A major problem with the twelve attributes of effectiveness as described by the CCAF is that with perhaps one exception, each is so broad that they are not likely to be useful for the development of standards. This point is crucial because the analogy to financial audits can be useful only if a set of standards equivalent to generally accepted accounting principles (GAAP) can be developed for program effectiveness.

The one attribute of effectiveness likely to be useful for developing standards is “financial results.” Standards such as GAAP can be set for the auditing of management representations of a program’s financial position because the units of measurement are agreed. Those units are units of currency—money. While there may be some imprecision in certain features of a financial statement such as the valuation of liabilities and assets, and

predictions of future performance of certain investments, the units to be measured are agreed across all programs.

That money is the agreed unit of measure is extremely important because auditing is basically a process of examining reports to determine the degree to which they comply with predetermined standards. If there is agreement that what is being examined is money, then the task of the auditor is only to examine whether it has been reported according to certain rules and whether the amounts reported can be substantiated from a more or less detailed examination of the records. Because money is the generally accepted variable of interest in financial reporting,

The role of the auditor is to add credibility to management's assertions on the financial position and results of the organization. This assurance is based on adherence to generally accepted accounting principles (GAAP). Management uses GAAP as an overall framework to guide their assertions. (CCAF, 1987;123)

However, as soon as audit is expanded beyond the "financial results" area, there are no structural equivalents to GAAP to guide the audit and it is no longer obvious that money is the appropriate variable to be measured. Thus, an auditor cannot give assurances that the management representations are relatively complete and accurate because there are no generally accepted standards to serve as the basis of such assurances nor are there any generally accepted standards regarding what should be measured. In other words, in the area of financial accountability, the auditor's task is to discover and report on the degree to which management has adhered to accepted standards in reporting about money. In all other areas, the auditor would have to develop standards as opposed to using those already generally accepted, and would also have to develop standards against which to assess whether the appropriate variables were measured and reported.

In the vast majority of public sector programs the most important intended results are not readily translated into units of money and the variables that should be measured and reported are not generally agreed. The examples of "Achievement of Intended Results" contained in *Effectiveness*... are important illustrations of this point.

In terms of solving major crimes in Peel, our officers had another very impressive year. In comparison to other major Canadian forces, our clearance rate ranked first in homicide, attempted murder, breaking and entering, and robbery. Put another way, no major police force in Canada outperformed the Peel Regional Police Force in areas where it matters most. (CCAF, 1987;92)

If you believe that the major role of the police is to enforce the law by apprehending and charging presumed law breakers, then clearance rates are important indicators of such aims. On the other hand, if you believe that the major role of the police is to keep the peace and prevent law violations, then clearance rates are of minor importance. Much more important are the incidences of crimes against persons and property. Similarly important are indicators of collaboration and respect between the police

and the rest of the citizenry, especially as reflected in rates of victim reporting and citizen assisted apprehensions.

Of the students who undertook continuous language training in 1982-83 under the revised Treasury Board policy of July 1981, 83% completed their courses within the maximum length of time prescribed. Only 2% did not succeed; the remainder either interrupted their training or succeeded with additional training. (CCAF, 1987;92)

Completion of training is clearly only one of a number of possible indicators of success of a language training program. Perhaps a better indicator would be the proportion of trainees who actually used their newly acquired language skills on the job. Or perhaps some direct measures of performance, independent from course completion, would provide better indicators of success. Others might argue that the important outcomes of such training are shifts in attitudes among trainees and that these should be measured by direct behavioral observations of those trainees in their post training job performance. Not only do they actually use their new language, but are they now more or less likely to communicate with those whose first language is the one they have just learned?

As can easily be seen, audit assurance that the summary statements quoted above were accurate reflections of the Peel Regional Police Force and of language training, respectively, would be analogous to an audit assurance of a financial report that said the arithmetic was correct but did not take any steps to verify the meaningfulness of the numbers being added. To know that a variable has been measured and reported accurately is meaningless if it is the wrong variable.

Evaluators have recognized for a long time that much of the debate in measuring effectiveness is appropriately centered on selecting which variables to measure as indicators of effectiveness. Once the variables have been selected, their adequate measurement is a relatively technical matter about which agreement is reasonably easy to achieve. It is the selection of the variables that generates the heat. It is also the selection of the variables and their subsequent operational definition that determines the meaning.

The selection of some variables as important and worthy of measurement and others as unimportant and not worthy of measurement is a value based and fundamentally political process. It is forming policy by determining how program implementation and the success of that implementation will be measured. Basically, the group that determines which indicators are indicators of success and which are indicators of failure controls the program. For auditors to enter this arena will require them to relinquish their mantle of disinterest and objectivity. Moreover, it will inevitably corrupt auditors in the sense that they will become stakeholders in the program. They will be required to defend not only their assurances that the facts reported by management are accurate, but also that management reported the right facts. In so doing they will become part of the accountability problem, not part of the solution.

At the beginning of this commentary the development of standards was described as a three phase process; selection of the attributes to be meas-

ured, agreement on how to measure them, and determining the range of performance acceptable on each attribute as measured. Clearly, the CCAF has completed none of these phases with regard to eleven of the twelve attributes in their definition of effectiveness. Since these attributes are much too broad to be measured and reported directly, decisions must be made about what variables should be measured and reported under each attribute. It is no accident that the systems for monitoring and reporting financial activities are more generally accepted as meaningful than systems for monitoring and reporting any other area of program activity. That is because money is generally accepted as an important thing to know about. There are not equally well accepted indices in any other area of program effectiveness.

This report offers no helpful suggestions about selecting variables. What it offers is a structural solution to a content problem. The structural solution offered by the CCAF depends upon the existence of a generally accepted set of variables. Unfortunately, by not addressing the issues related to the selection of variables to measure and report, the report appears to have set out an extremely naive position. Indeed, without a generally accepted set of variables, or a well articulated methodology for developing it, this report is "yet another list" (CCAF, 1987;82).

The variables selected for measurement directly reflect the values of the selectors and the theoretical frameworks and models that justify and conceptually organize the programs being measured. Significant progress has been made in the past decade in sensitizing program evaluators to the need for carefully selecting measures of effectiveness based upon program goals, program rationale, and stakeholder expectations (Bickman, 1987; Nowakowski, 1987). However, we are clearly not at the stage where these advances can be codified with any certainty. Alexander (1987) has argued persuasively that the selection of variables will often depend upon allegiance to frameworks that are not empirically testable. Conrad (1986) has presented convincingly the perils of selecting measures of effectiveness without carefully addressing the theoretical rationale of the program. To the extent that the CCAF report has essentially ignored these issues, it is neither a step forward nor a step backward, but a step to the side. Let us hope that this polite gesture will allow the rest of us to pass on quickly and continue to address the substantive issues of measuring and reporting program effectiveness.

References

- Alexander, Bruce K. (1987) The Disease and Adaptive Modes of Addiction; A Framework Evaluation. *The Journal of Drug Issues*, 17(1), 47-66.
- Bickman, Leonard (ed.), (1987) Using Program Theory in Evaluation. *New Directions for Program Evaluation*, no. 33. San Francisco: Jossey-Bass.
- Canadian Comprehensive Auditing Foundation (CCAF) (1987) *Effectiveness Reporting and Auditing in the Public Sector*. Ottawa, Ont.: The Canadian Comprehensive Auditing Foundation.
- Conrad, John (1986) The Assessment of Penological Effectiveness; The Rapsheet Model of Evaluation. *Canadian Journal of Program Evaluation*, 1(2), 57-67.

- Nowakowski, Jeri (ed.) (1987) *The Client Perspective on Evaluation*. *New Directions for Program Evaluation*, no. 36, San Francisco: Jossey Bass.
- Rossi, P.H. & Freeman, H.E. (1985) *Evaluation: A Systematic Approach (3rd Ed.)*. Beverly Hills, Ca.: Sage Publications Inc.