

DISCUSSION PAPER

**Addressing Attribution Through Contribution Analysis:
Using Performance Measures Sensibly**

John Mayne

Office of the Auditor General of Canada

June 1999

Contents

	Page
Introduction	1
Accountability for Outcomes	1
The Problem of Attribution	2
The Case of Performance Measurement	3
Recognizing the Limits of Measurement	5
Two Uses of Performance Measurement: Understanding and Reporting	6
Approaches to Attribution: Contribution Analysis	6
Doing the Best with Uncertainty	16
References	18

Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly

Introduction

A significant element of public sector reform in many jurisdictions is the move away from a management regime focussed on rules and procedures toward an approach that pays greater attention to the results being sought for citizens with taxpayers' dollars. Managing for results, results-based management and performance management have become common terms in public sector reform discussions (Auditor General of Canada 1997, Treasury Board Secretariat 1997, OECD 1997). The aim is to change the culture of public administration from one that is rules focussed to a culture focussing instead on the results that matter to citizens. This approach is characterized by measuring progress toward results that are sought, having the flexibility to be able to adjust operations to better meet these expectations, and reporting on the outcomes accomplished. Some jurisdictions have legislated this approach to public administration.

In many cases, progress has been made in moving in this direction. Nevertheless, the challenges of managing for results have been and remain significant, in particular the difficulty of measuring outcomes in the public sector in a cost-effective manner. Some of these problems are discussed below. There is an additional related problem that has not received enough attention: the need to rethink how we deal with accountability in this new management paradigm.

Accountability for Outcomes¹

In the past, accountability for the processes being followed, inputs used and perhaps outputs produced was most likely to be the regime in which public servants worked. This focus was consistent with the more traditional view of accountability focussing on what could be controlled and assigning blame when things go wrong. If the expected process was not followed, improper inputs were used or outputs were not delivered, then the responsible person could be identified and appropriate action taken, since one ought to be in control of the processes, the inputs used and the outputs produced. As such, there often was a reluctance to accept accountability for results beyond outputs, i.e. outcomes over which one does not have control. Being accountable for outputs has been much more

¹ This material is taken from a 1998 joint paper by the Office of Auditor General and the Treasury Board Secretariat.

acceptable to public servants than being accountable for outcomes. And in these cases, attribution is not a significant issue: it is clear that the program produced the outputs.

In the case of managing for results, and in particular outcomes, the degree of administrative control and scope for influence a federal manager has over the outcomes sought will vary considerably in different situations. In some cases, the federal program manager in question is the main player and has a quite significant degree of control over the outcomes. In other cases, the manager might be only one of several players trying, with the resources and authorities available, to influence the achievement of the intended outcomes. Effective accountability implies that managers understand these considerations, and have the means to deal with these more complex situations.

If the expected outcomes have not been accomplished, there may be several reasons, only one of which may be that the “responsible” manager hasn’t done a good job. The manager might have indeed done all that could be expected, but the results were not achieved due to circumstances beyond his or her influence. To encourage and support managing for results, we need a new view of accountability that acknowledges this more complex management world (Hatry 1997). Attribution here is a real problem.

Accountability for results or outcome ² asks if you have done everything possible with your authorities and resources to effect the achievement of the intended results and if you have learned from past experience what works and doesn’t work. Accounting for results of this kind means demonstrating that you have made a difference; that through your actions and efforts you have contributed to the results achieved. Finding credible ways to demonstrate this is essential if the move toward managing for results is to succeed.

The Problem of Attribution

Government programs are intended to produce certain outcomes: more jobs, a healthier public, better living conditions, etc. Effective programs are those that make a difference in meeting these kinds of objectives – they contribute to the intended outcomes that citizens value. In trying to measure the performance of a program, we face two problems. We can often—although frequently not without some difficulty—measure whether or not

² The terms outcomes and results are often used interchangeably. As used here, strictly speaking “results” includes outputs (see Figure 1) and hence is broader than outcomes. Nevertheless, much of the literature and some of the text here uses “results” to in fact mean outcomes, when the intention is clear. If a reference is being made to outputs, then “outputs” will be used.

these outcomes are actually occurring. The more difficult question is usually determining just what contribution the specific program in question made to the outcome. How much of the success (or failure) can we attribute to the program? What has been the contribution made by the program?

Despite the measurement difficulty, attribution is a problem that cannot be ignored when trying to assess the performance of government programs. Without an answer to this question, little can be said about the worth of the program; nor can advice be provided about future directions. Perhaps even without the program, the observed changes in outcomes would have occurred, or would have occurred at a lower level or later. In most cases, there are many other factors at play in addition to the impact of the program's activities. Such things as other government actions or programs, economic factors, social trends, and the like can all have an effect on outcomes. Managers, the government and taxpayers would like to know the program's contribution to assess the value of continuing with the program in its current form. Unless we can get some handle on this measurement problem, accountability for results will never take hold. The question is, how can we demonstrate that a program is making a difference

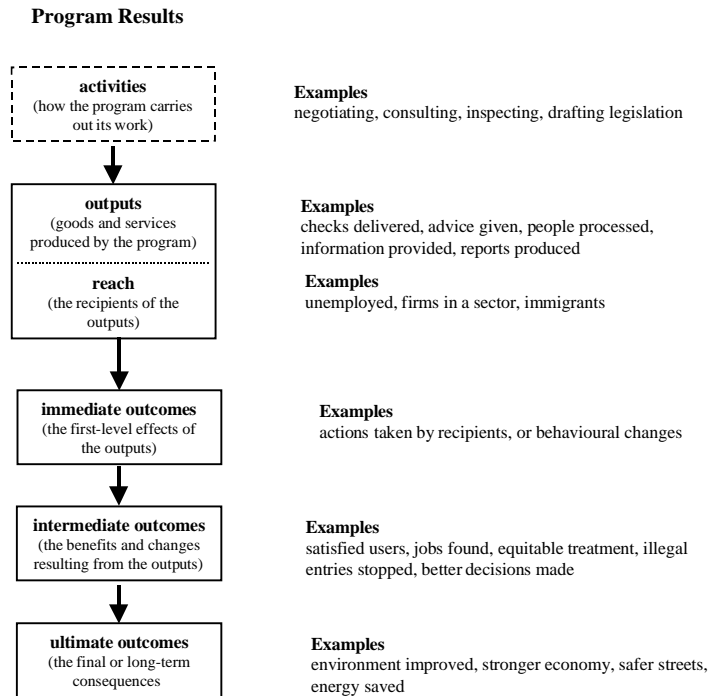
Policy and program evaluation is one measurement discipline that tries to provide answers to this attribution question³. Traditionally, it uses some form of controlled comparison to estimate what happens with the program in place versus what would happen without it. Extensive social science research methods have been designed with this problem of attribution in mind. And an evaluation study probably remains the best way to address this problem, if one has the time, money and expertise.

The Case of Performance Measurement

Performance measurement is extensively and increasingly used to measure the performance of government programs (Mayne and Zapico-Goni 1997). In comparison with evaluation, which usually undertakes special one-time measures and extensive analysis of the data gathered, performance measurement is characterized by regular and often more straightforward measurement of aspects of a program's performance. Performance indicators are used to track performance and feedback information to

³ The literature here is vast. See for example Hudson, Mayne and Thomlison (1992), Freeman and Rossi (1993) and Wholey, Hatry and Newcomer (1994).

managers and staff. They can form the basis for reports on what has been achieved by the program.



Performance measurement is often aimed at the very first level of impacts of a program, namely measuring the specific outputs (goods and services) provided by the program personnel. In these cases, the question of attribution is not likely to be a problem since there is an evident direct link between what the staff are doing and their immediate products. Increasingly, however, as we have seen, organizations are trying to measure or track the subsequent impacts of these services and products, the intermediate or even more final outcomes they are trying to accomplish. The attribution issue quickly surfaces. In the absence of a thorough evaluation study, what can be done

It is possible to structure a performance measurement system to directly try and get a measure of attribution. One could construct a careful time series and modify the program over time, tracking the resulting changes in all relevant factors. Or, in addition to measuring the impacts on those who are receiving the program, one could also measure the changes occurring in a similar comparison group that does not receive the program. To be successful, these approaches become in fact evaluations, using some form of quasi-experimental design.

While possible, this carefully constructed and often expensive measurement strategy is not usually associated with most performance measurement approaches. In the absence of an evaluation study, what can one do in the case of a “normal” or typical performance measurement or monitoring system to get a handle on the attribution issue? This is the question addressed in this paper.

Recognizing the Limits of Measurement

First we must recognize that determining definitively the extent to which a government program contributes to an particular outcome is usually not possible, even with a carefully designed evaluation study. We might be able to provide considerable evidence on a program’s impacts and might be able to significantly increase our understanding of how a program is impacting on a certain outcome, but in most cases of any complexity, there will not be a 100 percent guarantee. Rather, we need to talk of reducing the uncertainty in our knowledge about the contribution of the program. From a state of not really knowing anything about how a program is influencing a desired outcome, we might conclude with reasonable confidence that the program is indeed having an attributable impact; that it is indeed making a difference. We might also be able to provide a reasonable estimate of the magnitude of the impact.

Thus, we may need to rethink what measurement can usefully mean. Measurement in the public sector is less about precision and more about increasing understanding and knowledge. It is about increasing what we know about what works in an area and thereby reducing uncertainty.⁴ This view of measurement implies that we can almost always measure things, and in particular the contribution a program is making. That is, we can almost always gather additional data and information that will increase our understanding about a program and its impacts, even if we cannot “prove” things in an absolute sense. We need to include softer and qualitative measurement tools in our concept of measurement in the public sector.

The limits of measurement mean that we need to accept some uncertainty about the measures of performance we are likely to have available in many cases. If you must know with a high degree of certainty just what a program’s contribution is, then a well-designed evaluation is required. What we address in this paper applies in cases where one

⁴ For a discussion see the Auditor General of Canada 1996.

is willing or is required to make do with less certainty, where the aim of measurement is to acquire some insight and develop some comfort that the program is actually having an impact. This, we suggest, is or ought to be the aim of performance measurement. A good measurement strategy would include both ongoing performance measurement and periodic evaluation.

Two Uses of Performance Measurement: Understanding and Reporting

We need to distinguish two uses that can be made of performance measurement information. First, performance information can be used to better understand just what contribution a program is making. This is the management perspective, where one wants to use measurement to know more about if and how the program is making a difference; one is searching for knowledge. One wants to determine if the program is the appropriate policy tool to achieve the desired result. Here the question is how to use performance measurement as an investigative tool.

A second use of performance measurement is to explain or demonstrate the performance achieved by a program. In many jurisdictions, there is an increased focus and emphasis on reporting to Parliaments and the public what has been achieved with the tax dollars spent and authorities used. Performance measures frequently form the basis of such reporting. The question here is how can performance measurement information be used to credibly report on what has been accomplished; how can it be best used to report on the contribution being made by a program

We need to keep these two uses in mind as we consider how to deal with attribution using performance measures.

Approaches to Attribution: Contribution Analysis

What is needed for both understanding and reporting is a specific analysis undertaken to provide information on the contribution of a program to the outcomes it is trying to influence. Coupled with the comments above about the nature of measurement in the public sector, the task at hand might be best described as, for reporting, trying to paint a credible picture about the attribution of a program. For understanding, the task is to glean as much insight as possible from performance measures about how well the operations of the program are working. We suggest a number of strategies that can be used to address

attribution through performance measurement, as outlined in the box. Collectively, these are elements of a **contribution analysis**.

Contribution analysis attempts to explore and perhaps demonstrate what Hendricks (1996) calls “plausible association”; whether “a reasonable person, knowing what has occurred in the program and that the intended outcomes actually occurred, agrees that the program contributed to those outcomes?”

Acknowledge the problem. Too often, the measuring and particularly the reporting of performance through performance measurement systems completely ignores the attribution problem. The performance measured is either directly attributed to the program or attributed by implication, through the lack of any discussion or analysis of other factors at play. For anyone with even a little knowledge about the program and its environment, this kind of performance information will have little credibility. For managers, it provides no value-added information. In most cases, any number of factors can be advanced to explain the observed outcome other than the program itself. The more obvious these other factors are, the less credible is the performance information. Discussing other factors may also provide insight into the program itself, how it operated and its effects.

A first step then is simply acknowledging that there are other factors at play in addition to the program and that it is therefore usually not immediately clear what effect the program has had or is having in producing the outcome in question. Managers need to be realistic about the outcomes they are trying to influence if they want to gain new insight on how and if their activities are making a difference. For reporting, acknowledging the other factors at play is more honest and hence more credible than pretending they do not exist. As we will see below, there is more that can be done, but recognizing the other factors at play while still believing the program is making a contribution is a critical first step.

Analyze and present the logic of the program There is some logical reasoning behind the program that explains what it is supposed to be accomplishing and how. This logic or

Contribution Analysis: Addressing Attribution with Performance Measures

- Acknowledge the problem
- Present the logic of the program.
- Identify and document behavioural changes.
- Use discriminating indicators.
- Track performance over time.
- Discuss, and test alternative explanations.
- Gather additional relevant evidence.
- Gather multiple lines of evidence.
- When required, defer to the need for an evaluation.

theory might be quite convincing or well-established based on past experience. By developing the logical case, one can see what is supposed to or is believed to be happening. Constructing and presenting this theory of the program is a standard component of planning for an evaluation study (Wholey 1983), where often a logic chart is used (Julian, Jones and Devo 1995).

More recently, the power of this approach is increasingly seen in the performance measurement world where such terms as outcome sequence charts, results chains and “visible indicator tree” (Meekings 1995) are being used to describe the same diagnostic tool. In addition to getting a handle on the attribution issue, these tools are proving invaluable in designing and implementing performance measurement systems. Further, by forcing program designers to be clear about the problems that programs are designed to address and how to address them, logic models encourage programs to be more precise in their design.

A logic chart for a program tries to display on a page how the program is supposed to work—how, that is, the various outputs of the program are believed to produce a number of results that will lead to the intended final outcomes of the program. Logic charts can also discuss unintended impacts that might occur and need to be watched for, as well as the key external factors influencing outcomes.

Figure 1 illustrates in a generic fashion what a logic chart can look like; there are a variety of presenting one. A logic chart illustrates the linkages between specific outputs, specific intermediate outcomes and specific end outcomes. In others cases, it may be adequate to present a less complicated picture of the program logic. Figure 2 illustrates this case for an environmental program. Logic charts explicitly include the idea of reach—who the program is expected to reach—⁵ and immediate outcomes. This is because it is often at these levels that performance indicators can do a good job of measuring—that is, levels in the results chain over which the program typically has most control. Further, evidence that the intended immediate outcomes have in fact occurred is a critical step in demonstrating the larger performance story. In this manner, the program can be shown to have had some effect.

⁵ Montague (1998) discusses the importance of including reach into a logic chart.

Figure 1
A Program Logic Chart

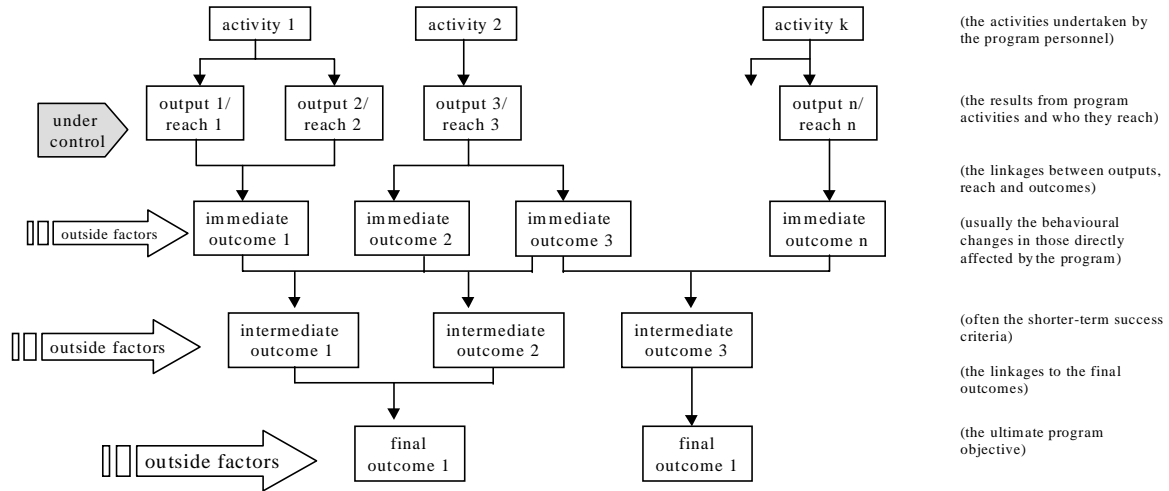
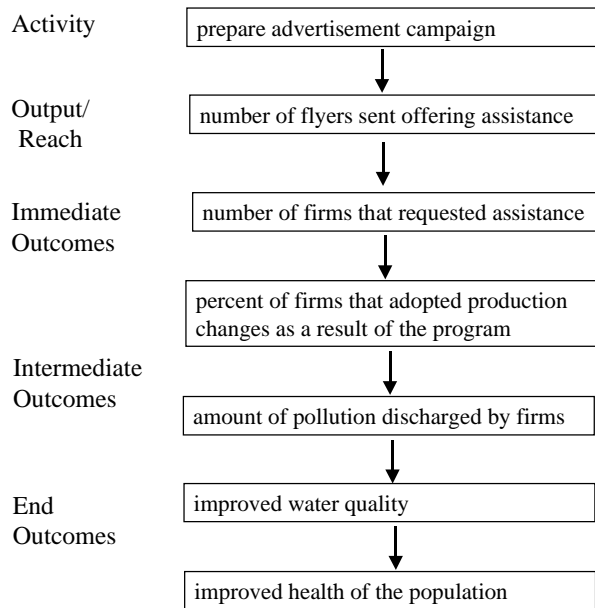


Figure 2
Logic Chart
Industrial Technical Assistance Program to
Reduce Water pollution (*)



(*) From Colombia: Paving the Way for a Results-Oriented Public Sector. Colombian National Department of Planning and the World Bank. January 25, 1996. p. 9.

Developing and using a logic chart has a number of benefits for program managers, such as developing consensus on what the program is trying to accomplish, developing an understanding on how it is believed to be working, clearly identifying the clients of the program, seeking and getting agreement on precisely what results are intended—the performance expectations—and identifying the key measures of performance. We are particularly interested in the additional benefits of identifying

- the cause-effect relationships implicit in the program’s theory;
- the outside factors at play; and
- areas where understanding about the impact of the program is weak.

Typically, some of the “links” between results in the logic chart are well known and have been established in past practice. There is less likely to be disagreement on their role in bringing about the intended impacts. Other links may not be so well accepted and those suggest where further evidence (i.e. additional performance measures) might be most fruitfully sought. Any additional evidence one can gather to confirm such links will add to understanding how the program is working and bolster the argument that the program is making a contribution. Similarly, if significant outside factors are identified as possibly having an effect on the intended outcome, then evidence to refute or determine the extent of influence of those claims will be useful in addressing the attribution question.⁶

In this way, managers can use the diagnostic tool of logic charts to better understand how they and others believe the program is working. They can design the program operations to fit these expectations. Through presenting and discussing the logic behind the program when reporting performance, one has laid out exactly what is being measured and what the major assumptions are concerning the contribution of the program. As a result, weaknesses in program assumptions are identified suggesting where more evidence is needed. At a minimum, reporting this way allows one to know what challenges to the credibility of the performance measures used can be raised.

⁶ In the case of reporting, we are not suggesting that only evidence that bolsters the claim of program impact should be gathered or sought. Being able to say with some confidence that it is not known what contribution the program is making is also valuable knowledge. We are trying to gather through performance measures as much evidence as is practical to understand the extent and nature of the contribution being made by the program and to support such a claim.

Identify, measure and document expected behavioural changes.⁷ In order to bring about an outcome, programs have to change people's behaviour. The outputs of the program must be aimed at influencing the program's clients or target audience – the reach element—to act in different ways so that the anticipated outcomes can occur. Logic charts often focus only on the sequence of events that are expected to occur and thus may be at too aggregate a level to detect the specific behavioural changes that must occur as prerequisites of each of the events. By trying to identify and then document the changes in attitudes, knowledge, perceptions and decisions taken by program target groups, which logically link to the outcomes being observed, a good understanding of the actual impact the program is having can often be acquired. Furthermore, these are often some of the immediate and intermediate outcomes that can be measured more readily. As a result, it may be useful to set performance expectations and targets at this level where there is a reasonable level of control (United States GAO 1998).

A more detailed logic chart is one approach, where the focus is on the specific behavioural changes resulting from the program's outputs that we can observe for those “reached” by the program. This requires clearly identifying who the various clients of the program are and how their behaviour is expected to change. If we can observe these short term changes occurring, the logical case for the program's attribution can be enhanced.

Thus managers either trying to better understand the effects of their programs or trying to report on performance can benefit from extending the analysis of logic charts to include consideration of the specific behavioural changes expected as a result of the program.

Use discriminating indicators. A good logic chart of a program often illustrates the many aspects of performance that could be measured and reported. Considerable care is needed in selecting indicators of performance. Here we are considering the attribution issue where it is important to use performance indicators that best discriminate or focus on the outcomes in question. Often, the indicators that are used relate only broadly to the circumstances of the program clients, the economy or society as a whole. With a little more thought given to how the program operates (from the analysis of the logic chart), the indicators can often be improved upon to more carefully focus on what specific

⁷ The ideas in this section were proposed by Steve Montague of the Performance Management Network, Ottawa.

benefits the program is intended to achieve. In particular, one can try and “refine the denominator” of the indicator.⁸

Many indicators are ratios, where the denominator qualifies the numerator. Consider a program designed to reduce air accidents by inspection of the air worthiness of aircraft. An indicator might be the number of air accidents per air-mile flown. A better indicator would be the number of air accidents due to structural failure per air-mile flown. But structural failures may occur regardless of inspections. Therefore, it may be better still to use two indicators: the number of air accidents per air-mile flown due to structural failure in aircraft inspected and the number of air accidents per air-mile flown due to structural failure in aircraft not inspected. By comparing structural failures in inspected and uninspected aircraft, one can estimate what inspection does to reduce the problems that inspection is designed to address. Questions of attribution still exist, but the more refined indicators reduce the problem and improve the chance of providing useful information on the contribution of the program.

Tracking performance over time or location. In cases where the program activities have varied over time, showing that outcomes have varied in a consistent manner with the variation in activities can strengthen the argument that the activities have indeed made a difference. In the simplest example, if an expected outcome has been observed after (and not before) the program activity has started up, this suggests the program is having an effect. In a more complicated case, if the outcome improves at sites (or at times) where the program has been implemented but not at others (such as a national program operating at many locations), the case for making a difference is even stronger. Hendricks (1996) identifies a number of such cases where by tracking performance measures we might show that:

- outcomes appeared at an appropriate time after our efforts began;
- outcomes appeared in different locations or with different people;
- outcomes faded when our efforts stopped;
- only those outcomes appeared that we should have affected;
- outcomes appeared only where or when we were active; and
- the biggest outcomes appeared where we did the most.

⁸ The term and the example were developed by Hugh McRoberts of the Office of Auditor General.

In some areas of programming, such as the impacts from research activities, there is likely to be a significant delay before the intended outcomes occur and the attribution picture portrayed through tracking performance over time will not be as evident. In these cases, one still needs to track outcomes over time to see if the intended outcomes have occurred, but demonstrating or understanding attribution is even more of a challenge. Some of the other approaches described in this paper need to be used.

Explore and discuss plausible alternative explanations. The attribution problem arises when one believes or is trying to claim that a program has resulted in certain outcomes and there are alternative plausible explanations. That is, those who are skeptical that it really was the program's contribution that counted will point to other reasons for the observed outcome—for example, other related government programs, economic or social trends, behaviour unaffected by the program.

Dealing with these alternative explanations explicitly is often the best way of buttressing an argument in favour of the program's impact. This entails:

- identifying the most likely alternative explanations;
- presenting whatever evidence or argument you have to discuss and, where appropriate, discounting these alternative explanations; and
- presenting whatever evidence there is that the program is a more likely explanation.

Of course, if there is little evidence that counters alternative plausible explanations, then you may have to conclude that you do not really know what the program's contribution has been and maybe (see below) suggest that an evaluation or further evidence is needed.

The kind of evidence that could be used to counter arguments for alternatives to the program depends on the program and its situation. But two generic types are available. First, there is a logic argument. One might refer to the theory behind the program and the kind of theory that would be needed to support claims for rival hypotheses. Supporting alternative explanations may involve assumptions more unlikely than those associated with the program. Second, one can bring actual evidence to bear concerning the alternative explanations, as discussed further on.

Addressing the attribution problem this way demonstrates that:

- you are aware of the complexity of the situation;
- you acknowledge and understand the other factors at play; and
- you are nevertheless concluding (assuming you are) that the most likely explanation for the observed outcome is that the program has made a significant contribution.

The burden of proof then falls on others to demonstrate that some other factor was the main factor in the chain of events that led to the outcome.

Unless you discuss alternative explanations, your claim about the program's efficacy can be effectively challenged by simply pointing out the existence of alternative explanations.

Gather additional relevant evidence. Performance measurement is about gathering evidence on the performance of a program. We suggest that some of that effort be devoted to evidence that would support statements about attribution.

As suggested earlier, one might gather evidence concerning alternative explanations of the observed outcome. This will mean gathering data such as contextual and historical information about the plausibility of the alternative explanations. The data might be part of the routine performance measurement system, but more likely would be collected from time to time when analysis of the program's contribution is undertaken. Data collection might entail a review of the relevant literature, surveys, tracking of relevant external factors, field visits, or focus groups. The stronger the case that can be made, the stronger is the conclusion about the program's contribution.

In addition, one might try and gather evidence about the contribution of the program directly, most often through the use of **expert opinion**. In many program situations, there are persons outside the program who are seen as knowledgeable about the program area, the program's impacts and the environment in which the program operates. A structured survey may be able to provide some evidence, albeit subjective in nature, of the extent to which the program is influencing an outcome. Surveying such individuals is often done to find out other information about the program, in which case adding questions on attribution is not very expensive. A focus group of experts may be another approach that would allow some probing as to why views are held. In the absence of other more costly

data, this approach can be a relatively inexpensive way to increase comfort about the influence of the program.⁹

Two other sources of data are often overlooked. There is frequently considerable existing data available from **program files**, some of which might be useful to provide information on the contribution of the program. This type of existing data, which probably has been collected for other purposes, can often contain valuable information, particularly if used in conjunction with new data collected. In other cases, there may be useful **secondary analysis** available—studies that others have done in the program area that might clarify measurement and attribution issues. In still other cases, there may be **meta analysis** that has been done—analysis that synthesizes a number of studies in an area.

Finally, use can often be made of **case study evidence** on a program's outcomes. In programs where there are specific cases, projects or events, the evidence on attribution on one or two of these can be quite compelling; it can reveal the real nature of the program and also demonstrate, at least in these cases, that one can be fairly confident about the impact of the program's activities. In addition, case studies can also illustrate whether the program logic is indeed logical and reasonable (or not). This type of evidence can be quite persuasive but appropriate cautions are a must, especially when it is quite anecdotal. Case study and anecdotal evidence is best when illustrating a concrete case to complement other evidence that has been collected. On its own, however, it can be quite misleading since it may merely be one of the few cases that appears to have worked while the vast majority have not, as the US GAO (1996) recently found in a review of “Success Stories” of the US Department of the Environment. Further, there is a temptation for readers to generalize from anecdotal evidence, which should be cautioned against. Nevertheless, if the context and limitations are made clear, there is often a useful role for individual case studies.

Use multiple lines of evidence. We have discussed a number of ways to deal with the attribution problem. We suggest that the more ways that are used in any one case, the more definitive information we will have on attribution. This is the “multiple lines of evidence” argument. While no one piece of evidence may be very convincing, a larger set of different and complementary evidence can become quite convincing. Thus, in trying to

⁹ One caveat here is that if an individual expert has a vested interest in the program, then his or her views will need to be suitably discounted

reduce the uncertainty surrounding attribution, using as many lines of evidence as possible is a sensible, practical and credible strategy.

Defer to the need for an evaluation. In some cases, if the various lines of evidence point in different directions, there may be little one can say with enough credibility about the contribution of the program. If it is critical to have good information on attribution, then the best strategy may be to simply acknowledge that one does not know and suggest that an evaluation be carried out to address the attribution question. In most cases, however, if the program has indeed made a significant contribution, the various lines of evidence will confirm this.

Doing the Best with Uncertainty

We have argued here that what is needed in dealing with attribution using performance measurement information is to explore the issue in a systematic way and, when reporting, to paint a credible picture of attribution to increase our knowledge about the contribution being made by the program. We need to accept the fact that what we are doing is measuring with the aim of reducing the uncertainty about the contribution made, not proving the contribution made.

We suggest undertaking a **contribution analysis** that would examine and present the best case possible – a credible performance story - for attribution with the available evidence.

A credible performance story. Using contribution analysis, a reasonable case that a program has indeed made a difference would entail ¹⁰

- well-articulated presentation of the context of the program and its general aims;
- presentation of plausible program theory leading to the overall aims. (The logic of the program has not been disproven, i.e. there is little or no contradictory evidence and the underlying assumptions appear to remain valid;
- highlighting the contribution analysis indicating there is an association between what the program has done and the outcomes observed; and
- pointing out that the main alternative explanations for the outcomes occurring, such as other related programs or external factors, have been ruled out or clearly have only had a limited influence.

¹⁰ Hendricks (1996) proposes a similar list.

If all this is not enough, and there are too many gaps in the story, one ought to admit it and accept the need for an evaluation to provide better understanding of the contribution of the program.

Recognition of the problem and an understanding of the other factors at play will likely lead to additional data and information gathering. The result will be a better understanding of the program and how it is expected to work, and perhaps a redesigning of the program to reflect this enhanced understanding. In addition, better performance information will provide for a more credible demonstration of the impacts of the program through performance measurement.

References

- Auditor General of Canada (1996). *Report to the House of Commons: Matters of Special Importance*. Ottawa. p. 21.
- Auditor General of Canada (1997). *Report to the House of Commons: Chapter 11, Moving Towards Managing for Results*. Ottawa.
- Freeman, H. E., and Rossi, P. (1993). *Evaluation: A Systemic Approach*. Sage.
- Hatry, Harry (1997). We Need a New Concept of Accountability. *The Public Manager*. 26(3): 37-38.
- Hendricks, Michael (1995). *Performance Monitoring*. Presented at the International Evaluation Conference, Vancouver. November 4.
- Hendricks, Michael (1996). *Performance Monitoring: How to Measure Effectively the Results of Our Efforts*. Presented at the American Evaluation Association Annual Conference, Atlanta. November 6.
- Hudson, J., J. Mayne, and R. Thomlison, Eds. (1992). *Action-Oriented Evaluation: Canadian Practices*. Wall & Emerson.
- Julian, D. A., Jones, A., and Devo, D. (1995). "Open Systems Evaluation and the Logic Model: Program Planning and Evaluation Tools." *Evaluation and Program Planning*. 18(4): 333-341.
- Mayne, J. and E. Zapico, Ed. (1997). *Performance Monitoring for Public Sector Reform: Future Directions from International Experience*. Transaction Publishers.
- Meekings, A. (1995). "Unlocking the Potential of Performance Measurement: A Practical Implementation Guide." *Public Money & Management*. October-December: 5-12.
- Montague, S. (1998). *Build Reach into Your Logic Chart*.
<http://www.pmn.net/contributions/reachlog.htm>
- Office of the Auditor General of Canada and Treasury Board Secretariat (1998). *Modernizing Accountability Practices In the Public Sector*.
- OECD (1997). *In Search of Results: Performance Management Practices*. Paris.
- Wholey, J. S. (1983). *Evaluation and Effective Public Management*. Little, Brown and Co.
- Wholey, J., Hatry, H., & Newcomer, K. (1994). *Handbook of Practical Program Evaluation*. Jossey-Bass.
- United States General Accounting Office (1996). *Observations on DOE's Success Stories Report*. Testimony before the Subcommittee on Energy and Environment, Committee on Science, House of Representatives (GAO/T-RCED-96-133).
- United States General Accounting Office (1998). *Managing for Results: Measuring Program Results That Are Under Limited Federal Control*. (GAO/GGD-99-16).