

The New Policy on Evaluation: Straightjacket or Platform?

Presentation to the National Capital Chapter
Canadian Evaluation Society

Greg Mason
PRA Inc.
January 28, 2010

The Policy on Evaluation

Comprises three elements:

- Policy on Evaluation
- Directive on the Evaluation Function
- Standard on Evaluation for the Government of Canada

The Policy is arguably the most important document, but the other two should not be ignored.

One sub-theme, not apparent, is that the Policy on Evaluation attempts to raise evaluation to a similar level of important.

Policy on Evaluation

Main Goal

1. Defines the obligation for departmental evaluation plans to demonstrate progress toward achieving coverage of direct program spending over five years
2. Plans that do not demonstrate evaluation coverage of all direct program spending need to use a “risk-based” approach to planning coverage

The evaluation plan needs to either show 100% coverage or identify the programs that will be assessed (and not assessed) within the 5-year cycle using a **risk-based criteria**.

Risk Assessment: Materiality

- **Proportion of budget**

Bob Woodward: The story is dry. All we've got are pieces. We can't seem to figure out what the puzzle is supposed to look like. John Mitchell resigns as the head of CREEP, and says that he wants to spend more time with his family. I mean, it sounds like bullshit, we don't exactly believe that...

Deep Throat: No, heh, but it's touching. Forget the myths the media's created about the White House. The truth is, these are not very bright guys, and things got out of hand.

Bob Woodward: Hunt's come in from the cold. Supposedly he's got a lawyer with \$25,000 in a brown paper bag.

Deep Throat: Follow the money.

All the President's Men

Risk Assessment: Strategic Position

- Programs that are integral to the Department Agency priorities
- Programs, that if they fail:
 - Compromise the delivery of other programs
 - Compromise the priorities of the department
 - Potentially could cause social, economic, and political cost disproportionate to their magnitude
- Programs that are constitutionally and legislatively required can be assigned a lower priority since the discretion on spending is lower.
 - Note that the Policy identifies these as requiring only an “administrative review.”

Observation

- The Evaluation Policy appears to mandate 100% coverage, but:
 - Evaluation heads and deputy heads can develop a case for a risk-based approach that sets aside/defers some elements.
 - Such a case requires careful planning
 - The evaluation plan that does not present 100% coverage over the 5 year cycle, must present a credible argument that the “omission” or “deferral” does not compromise reporting on results or the Standard on Evaluation.
- This is not necessarily easy or even possible in some cases, but that option does exist and should be explored.
- The case for excluding programs on the grounds of low risk (to the Departmental Mandate and to the Standard on Evaluation) should be done early, and not in response to shortage of evaluation funds later in the five year cycle (credibility).

Core issues

- Reduced from four theme areas
 - Rationale/relevance
 - Design/delivery
 - Success/impacts
 - Cost effectiveness/alternatives
- To two
 - Relevance
 - Performance
- This appears to be a retreat, but the two core issues offer all the needed scope.

Relevance

Core Issues	
Relevance	
Issue #1: Continued Need for program	<i>Assessment of the extent to which the program continues to address a demonstrable need and is responsive to the needs of Canadians</i>
Issue #2: Alignment with Government Priorities	<i>Assessment of the linkages between program objectives and (i) federal government priorities and (ii) departmental strategic outcomes</i>
Issue #3: Alignment with Federal Roles and Responsibilities	<i>Assessment of the role and responsibilities for the federal government in delivering the program</i>

- #1 Implied in the term *demonstrable need* is whether “private sector” opportunities exist (or have been displaced) or whether other orders of government may be better positioned.
- #2 Horizontal initiatives challenge easy responses to this otherwise straightforward issue.
- #3 A key issue in roles and responsibilities is federal jurisdiction and constitutional alignment.

Performance

Performance (effectiveness, efficiency and economy)	
Issue #4: Achievement of Expected Outcomes	<i>Assessment of progress toward expected outcomes (incl. immediate, intermediate and ultimate outcomes) with reference to performance targets and program reach, program design, including the linkage and contribution of outputs to outcomes</i>
Issue #5: Demonstration of Efficiency and Economy	<i>Assessment of resource utilization in relation to the production of outputs and progress toward expected outcomes</i>

- #4 – We see a blurring of the old formative/summative evaluation. Also apparent is the need to show a causal/attribution link and a validation of program theory. These are all welcome specifications.
- #5 – This is the cost effectiveness issue linked specifically to **economy** (are we acquiring resources/inputs at the lowest cost) and **efficiency** (are the outputs being produced at the lowest unit cost). The **effectiveness** question (cost per unit outcome) is implied in the term “progress toward expected outcomes.”

The evaluation policy – straightjacket or platform

- There is nothing inherent in the policy that confines evaluation.
- It offers general direction and the possibility that evaluation managers and deputy heads can fashion the effective evaluation programs.
- The real issues revolve around:
 - Collateral support to the evaluation function
 - The context of the evaluation workforce
 - Uncertainty as to how to manage the 100% evaluation requirement.

Collateral Support to Evaluation

- Treasury Board has defined the requirements, but the guidance is general.
- Treasury Board is planning to re-issue the guidebook for evaluation, which will include detailed and concrete methodologies with examples of best practice.
- In the meantime, the absence of this guidance has created uncertainty and to a degree, some resentment.
- For example, more concrete information is needed on “the risk-based approach to planning coverage of direct program spending” to allow deputy heads and managers to identify those programs, whose evaluations can be deferred.
- Two specific requirements to improve support:
 - All TB analysts must share a common understanding of how Section 1.3.1b. may be applied.
 - TB must bring program management into the evaluation loop and insist on active support for the evaluation function, in the same way that audit receives support.

The context of the evaluation workforce

- The federal workforce is going through a massive recycling
- Corporate memory is being lost and a large fraction of the evaluation workforce is just starting their careers.
- Potential remedies:
 - Despite the rule that previous civil servants not receive contracts, senior program managers must be retained after retirement to provide key context, especially for relevance issues.
 - Training, training, training
 - Key expertise lies in the consulting community and creating executive “secondment” may allow that expertise to be shared without “raiding” the private sector.
 - Time

Uncertainty exists on how to manage the 100% evaluation requirement

Aside from the need to identify and make the case for what programs may be deferred, the primary issue is:

Failure to adapt evaluation methods to the new policy

Managing within the 100% requirement

- Clustering
- Undertaking improved evaluation planning
- Rationalizing evaluation methodology

Clustering

- Many speak of “cluster evaluations” where “cognate” programs are grouped, often by moving up the organizational chart.
- This will usually produce superficial reports
- An alternative is to create a *departmental logic model* using the RPP and DPR as a guide to **group the evaluations according to outcome** (intermediate)
- Defer evaluations of programs (expenditures) that make no material contribution to the outcome(s) in question (under section 1.3.1b).
- Note; the intent of the new policy is to have 100% coverage, and so this section only supports a deferral, not an omission.

Undertaking improved evaluation planning: Rehabilitating the RMAF

The RMAF should

- answer Relevance completely
- provide program context (theory of change, comparison with other jurisdictions...).
- present a detailed evaluation plan
- be the “contract” to deliver performance data for the evaluation (signed by program managers, evaluation managers, deputy head and TB)

Programs that cannot pass the relevance test should move to strategic review and not be evaluated.

If the data do not exist to support performance measurement, then the program should not be evaluated.

“As we know, there are known knowns; there are things we know we know.

We also know there are known unknowns; that is to say we know there are some things we do not know.

But there are also unknown unknowns -- the ones we don't know we don't know.”

Donald Rumsfeld

Rationalizing evaluation methodology

- Much of evaluation methodology has become “rote” comprising:
 - Document review
 - Literature review
 - Interviews
 - Case studies
 - Surveys (clients, managers, stakeholders....)
 - ...
- The methods are often completed serially, with a discrete report at each stage.



PRA Inc.
Research & Consulting
www.pra.ca
admin@pra.ca

Critique of current methodology

Document reviews

- Are not a method with any purpose other than to complete the program profile.
- The program profile is not possible to complete without some interviews to offer context.
- Therefore; the goal of the document review is to complete a program profile which really is a main element of the RMAF
- At most, evaluations should review recent documents (finances) to update the RMAF and evaluation assessment.



PRA Inc.
Research & Consulting
www.pra.ca
admin@pra.ca

Critique of current methodology

Literature reviews

- Critical to the assessment of the theory of change and program design.
- Ideally, these require peer and expert review to be credible
- Positioned within many evaluations, they are often small scale, partial, biased, (selective) and inadvertently support the program uncritically.
- Literature reviews are best positioned as part of strategic review and take a broader view of policy.



PRA Inc.
Research & Consulting
www.pra.ca
admin@pra.ca

Critique of current methodology

Interviews

- Have become a primary line of evidence, usually in the absence of performance data
- Many evaluations treat this inherently qualitative information as “countables”, under the argument that all interviewees are equal.
- In fact, while interviewing is often necessary to manage stakeholder expectations and can often yield good information on operational issues, a survey (mail/web) is often a superior method
- Key information interviews have limited value in measuring performance - the performance data should speak clearly
- Interviews, with senior program management, serve best at the program relevance stage (RMAF) and at the final evaluation to explain program operations/design and to help frame constructive recommendations for change. Very few evaluation would need more than 10 interviews.



PRA Inc.
Research & Consulting
www.pra.ca
admin@pra.ca

Critique of current methodology

Case studies

- Increasingly used to show performance for G&Cs programs, but
 - They tend to be biased (positively)
 - Are evocative and can dominate the evaluation
 - Often feature interviews with project proponents and use the self reports of project proponents as the “data”
- One approach to increase validity and economize on evaluation resources is to use internal audits (or external audits) on funding to third parties.
 - Augmented by reports from the funded recipients, these audits can increase the reliability and validity of case study material

Conclusion

- Evaluations are often completed at the end of the five-year funding cycle
- They are rushed and superficial
- The evaluation process should be seen as a three- stage effort over the program cycle:
 1. RMAF to respond to relevance with
 - An assessment of performance data and
 - A embedded contract to deliver performance data signed by program and evaluation management
 - Regular reviews of progress on performance reporting
 2. Evaluation assessment to finalized the design
 3. The formal evaluation

Conclusion

- Nothing in the policy, standard or directive confines evaluation
- The 100% rule is not an unreasonable expectation of the taxpayer
- More guidance is needed from TB and program managers need to be drawn into the evaluation process
- Effective planning, evaluation over the program cycle and an adjustment of evaluation methods all contribute to meeting the goals of the policy.