Research and Practice Notes / Notes sur la recherche et les méthodes

Finding Balance: An Evaluation Governance Model to Ease Tension between Independence and Inclusion

Robert K. D. McLean, Kaitlyn Finner, and Lisa Woodward
International Development Research Centre

François Dumaine
PRA Inc.

Kathryn Graham
Alberta Innovates

Jocelyn E. Mackie
Grand Challenges Canada

David Peckham
Canadian Institutes of Health Research

Kathryn Wehr
Robert Wood Johnson Foundation

Abstract: Practitioners and theorists have documented the benefits of user engagement and participation in evaluation and, at the same time, the value of neutral and impartial evaluative evidence. Yet producing both an independent and inclusive evaluation is a leading challenge in our field. In this practice note, we present one solution. We describe the design of an evaluation governance structure that was used to find balance between these two themes. We also identify key elements of this experience and present these for adaptation by others, given appropriate tailoring. We have documented our experience as we believe that governance provides currently uncharted potential for this discipline spanning challenge.

Keywords: evaluation governance, evaluation management, independence, oversight, participation

Résumé : Les spécialistes et les théoriciens ont documenté les avantages de la participation des utilisateurs en évaluation et, de la même façon, la valeur de
Données d’évaluation neutres et impartiales. Il reste que réaliser une évaluation aussi indépendante qu’inclusive est encore un défi de taille dans notre domaine. Dans la présente note de pratique, nous présentons une solution. Nous décrivons le design d’une structure d’évaluation de la gouvernance qui a servi à trouver un équilibre entre les deux thèmes. Nous nommons aussi les éléments clés de cette expérience et expliquons comment ils peuvent être adaptés pour répondre à des besoins particuliers. Nous avons documenté notre expérience puisque nous croyons que la gouvernance est une avenue prometteuse pour relever ce défi qui touche de nombreuses disciplines.

Mots clés : gouvernance en évaluation, gestion de l’évaluation, indépendence, surveillance, participation

INTRODUCTION

This practice note aims to detail the management and oversight process that was used to govern the 2015 summative evaluation of the Development Innovation Fund—Health (DIF-H), a consortium effort of Canada’s International Development Research Centre (IDRC), Grand Challenges Canada (GCC), and the Canadian Institutes of Health Research (CIHR). We believe it will be of interest to evaluators, evaluation managers and commissioners, and organizational leaders who are interested in governance models for evaluation. Stakeholders require timely and actionable results for implementation, particularly where independence and accountability are important objectives of evaluation. As a practice note, this manuscript describes a record of experience that we present for critique, debate, or inspiration. Its purpose is not to provide generalizable research evidence.

Readers of this journal will know that producing independent and useful evaluations is an ongoing challenge that requires continuous attention. We have co-crafted this practice note with two objectives in mind. First, it is the output of an after-action review process to advance our own learning and improvement agendas (National Aeronautics and Space Administration, 2016; Overseas Development Institute, 2009). Second, it is meant to ensure that what we perceive as a successful attempt at addressing this challenge is documented and shared with others who may find themselves in similar circumstances.

The desire to share this experience was also driven by the fact that the evaluation of the DIF-H, while involving a federal agency and a Crown corporation of the Government of Canada, was not required to be undertaken using the approach outlined in the Treasury Board’s Policy on Evaluation (now the Policy on Results). This provided some leeway in exploring new approaches to managing evaluations, while also considering some of the fundamental principles upon which federal evaluations rest.

This practice note begins with an introduction to the DIF-H (the object of evaluation) and the context in which the evaluation took place. Next, we introduce the concepts of independence and inclusion as central themes and challenges to evaluation management, highlighting our reasons for attempting to address...
these themes through internal-external governance. The third section outlines the internal-external governance approach and how it was implemented. Finally, we provide self-reflection from all members of the governance committees. We note what we thought worked well and areas we would improve in a second go. We conclude the practice note by offering some remarks to others who may also consider applying this approach, with appropriate adaptation.

**The Development Innovation Fund—Health**

The Government of Canada established the DIF-H in 2008. As announced in that year’s federal budget, the fund aimed to “support the best minds in the world as they search for breakthroughs in global health and other areas that have the potential to bring about enduring changes in the lives of millions of people in poor countries” (Government of Canada, 2008). With a budget of $225 million over seven years, the DIF-H implements this complex mission by providing financial, and other forms of, support to competitively selected proposals from Canadian and low- and middle-income country researchers who present bold and innovative ideas that address humanity’s most difficult health challenges.

A three-party consortium implemented the DIF-H. It includes IDRC, the CIHR, and GCC. Each plays a discrete role. GCC is a purpose-built, not-for-profit organization directed by the Government of Canada to be the DIF-H implementation arm, responsible for the delivery of the identified policy priorities and the monitoring and evaluation of the funded projects. The CIHR, Canada’s national health research funding agency, is responsible for organizing the competitive scientific peer review of applications to the fund. IDRC, a Crown corporation of the Government of Canada, is responsible for advising the Canadian government on the progress of the DIF-H, disbursing funds to GCC, and managing the evaluation and oversight of the DIF-H as a whole (see Figure 1). Since it is not the purpose of this practice note to describe the DIF-H or the concept of Grand Challenges for Development, we encourage interested readers to review the website of GCC and IDRC to learn more about the DIF-H, and the Grand Challenges for Development website to learn more about the global Grand Challenges effort.

![Figure 1. The DIF-H Consortium Model](image-url)
The context of the evaluation

Responsibility for the management of the evaluation of the DIF-H was directed to IDRC by the Treasury Board Secretariat of Canada through a memorandum of Cabinet. A specific requirement was that a summative evaluation would be completed to provide the Government of Canada with a view of the DIF-H’s performance and results. Early meetings with senior government officials stressed the importance of rigour and strict independence of the evaluation. As such, IDRC concluded that the evaluation should be commissioned by an external evaluator, based outside of Canada, with no prior affiliation to the DIF-H or any of its consortium members, with a reputation for evaluation excellence, in addition to expertise in innovation and health.

At the same time, since IDRC, GCC, and CIHR follow closely the concept of evidence-based decision making, they understand that the apt use of evidence is time dependent (Pawson 2006). In this context, waiting for external recommendations for action was not ideal. As such, during the inter-consortium evaluation planning meetings, a desire became clear that we wanted to do all we could to ensure that the themes of independence and summative judgment that were now driving the evaluation would not undermine the production of accessible and actionable evidence. Moreover, we also wanted to make sure that this evidence would reach the consortium members in relatively real time and in the right learning spaces (for example, appropriate levels of management, appropriate decision-making moments, and so on). Thus was born our aspirational goal to weave the themes of independence (for accountability) and inclusiveness (for utility and real-time use) through the fabric of the evaluation process. The final DIF-H evaluation report can be viewed in the Open Access Digital Library of IDRC and is available on the webpage of GCC (Oxford Policy Management, 2015).¹

DESCRIPTION OF EVALUATION MANAGEMENT FOR INDEPENDENCE AND INCLUSIVENESS

At a minimum, an independent evaluation is one in which the evaluator(s) “are not beholden to those who designed and implemented the intervention” (Morra Imas & Rist, 2009, p. 32; Organisation for Economic Co-operation and Development, 2002). Bias or partiality on behalf of the evaluator risks undermining the credibility, reliability, quality assurance, and ethics of his or her work regardless of whether the lack of independence is real or merely perceived (Picciotto, 2008).

Yet, independence is not the only factor in evaluation, as was the case for the DIF-H, where utility and relevance to stakeholders were also important. In fact, the concept of independence can seem to be at odds with utility as utilization-focused evaluations encourage the evaluator to engage with evaluation users—often the very people that have designed and implemented the intervention—to better ensure a usable and relevant evaluation (International Development Research Centre, 2012). The inclusion of end-users and their considerations throughout the evaluation is premised on the idea that the involvement of primary
users is central to producing an evaluation conducive to learning and the application of that learning (International Development Research Centre, 2012; Patton, 2008). This inclusion can result in a (real or perceived) tension for evaluators as they must navigate between the safeguarding of their independence and the development of a useful and actionable end product. We contended, however, that a balance could be facilitated between independence and inclusiveness if an appropriate governance structure was established to act as a fulcrum (see Figure 2).

Despite the wealth of literature dedicated to the broad concept of management (Compton, 2009), there is a lack of work focused on the management of evaluations (Baizerman & Compton, 2009; Compton, 2009). This literature gap presents both a challenge and an opportunity. On the one hand, there is little experience to draw on to inform governance structures for managing evaluations. On the other hand, such a situation demands creativity.

The literature on advisory groups is a starting point in understanding how governance procedures can play a role in evaluation management. An advisory group’s role can differ depending on the scope and nature of an evaluation, but they are generally established to provide feedback throughout the evaluation process, based on their expertise (Baizerman, Fink, & Roholt, 2012; Compton, Baizerman, & Roholt, 2010). These contributions can include social and political insights to assist in understanding context; however, these insights are non-binding as advisory groups generally do not have governing authority over evaluations (Baizerman, Fink, & Roholt, 2012). Advisory groups can therefore bring fundamental contextual considerations and subject area expertise to an evaluation.

Similar to advisory groups, steering committees can be used in evaluation as part of the governance structure. Most common in multi-partner evaluations of international cooperation initiatives, steering committees have “a broad membership of all interested partners” and are characterized by a complex group of stakeholders with divergent interests (Dabelstein & Kliest, 2013, p. 42). For the

Figure 2. Governance as a Balancing Fulcrum
DIF-H, IDRC desired the contextual knowledge and input of consortium members to guide the evaluation toward an accurate and useful result. At the same time, IDRC could not tolerate interference or the perception of interference in the evaluation (including from IDRC itself) as full external independence of the evaluation was a formal requirement of the Government of Canada. Since no model was immediately available to us in the literature or practice record, IDRC devised the approach to governance using a dual committee, as described below.

OVERVIEW OF THE PRACTICE IMPLEMENTED: DUAL COMMITTEE (INTERNAL-EXTERNAL) GOVERNANCE

To practically address the challenge of blending independence and inclusiveness into the DIF-H summative evaluation, IDRC developed an evaluation management plan that embedded both themes directly into the governance structure of the evaluation. It is important to note that this management plan was clearly detailed and socialized with the consortium members and with the primary evaluation user, the Government of Canada (International Development Research Centre, 2015), all of whom supported the approach. Since IDRC was ultimately responsible for the DIF-H evaluation, the governance structure was developed with the purpose of supporting IDRC in its evaluation management role. The goal was to ensure that any oversight upheld an open process and also ensured that the independent evaluator was not “captured” by the consortium or any of its members (Laffont & Tirole, 1991).

To achieve this, the governance structure comprised two discrete committees, with distinct roles and responsibilities. IDRC labelled these two groups: the DIF-H Evaluation Committee and the DIF-H External Oversight Committee. In our implementation, the two committees did not meet directly with each other. Instead, information was shared via recorded meeting minutes and formal records of decision.

The DIF-H Evaluation Committee

The DIF-H Evaluation Committee was charged with the regular discussion of issues pertinent to the evaluation. Chaired by the Evaluation Office of IDRC, the committee was composed of senior level representatives of each consortium member. In plain terms, the membership of this committee had a vested interest in the DIF-H. The terms of reference of the committee outlined that it would meet at a minimum of four critical check-in points:

1. at the evaluation initiation to discuss the management approach and committee roles and responsibilities;
2. at the evaluation design stage to discuss the evaluation approach presented by the external evaluator;
3. at the findings stage to react and respond first hand to evaluation findings before formal public reporting; and
4. at the evaluation conclusion to contribute to the management response and action plan, which were written by the consortium to address the recommendations of the evaluation.

However, the committee met more often than this, as it served as an ad hoc meeting place for discussions regarding the progress of the evaluation and the related business required of each consortium member—for example, data access for the evaluator and the scheduling of data collections activities. As such, it allowed for appropriate engagement between the necessarily external evaluator and the consortium members in an open and transparent way, as all consortium members were to be present for interactions (see Table 1). The DIF-H Evaluation Committee also created a learning and discussion space for consortia members. Via the committee, the process results and early findings of the evaluation were shared with each agency in a timely manner. This timing allowed for an early discussion of the emergent issues and for agency implementation or reaction to information coming from the evaluation in relatively real time. Nevertheless, at all times, the committee remained a discussion forum, playing an advisory role to the independent evaluator and the External Oversight Committee.

**The External Oversight Committee**

The External Oversight Committee was charged with decision making by the process of consensus on critical evaluation issues. It comprised three external subject area experts, who were deemed to be independent of any consortium member. To elect members to the committee, the President/Chief Executive Officer of each consortium member nominated an external representative, although each agency was granted veto rights on any other agency’s nomination. Once elected, committee members were asked to sign a declaration of their independence from the DIF-H or any consortium member, including their nominating agency. This process ensured that no evaluation “decision maker” on this committee held a vested interest in a positive or negative outcome of the evaluation. The committee was chaired by a non-voting representative of the IDRC’s Evaluation Office (see Table 2).

The terms of reference dictated that the committee would be convened at a minimum of three decision-making points:

<table>
<thead>
<tr>
<th>Table 1. The DIF-H Evaluation Committee</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chair</strong> – Head of Evaluation, IDRC</td>
</tr>
<tr>
<td><strong>Membership:</strong></td>
</tr>
<tr>
<td>• CIHR – Chief Audit and Evaluation Executive</td>
</tr>
<tr>
<td>• GCC – Vice-President Operations and General Counsel</td>
</tr>
<tr>
<td>• IDRC – Senior Program Specialist in Evaluation</td>
</tr>
</tbody>
</table>
1. to select the evaluator, by reviewing applications and applying their expert judgment against the criteria of evaluator ability for the job as well as feasibility and rigour of the proposed approach submitted within the evaluation;

2. to approve the evaluation design, by applying their expert judgment in examining the evaluator’s detailed methodology report and the DIF-H Evaluation Committee’s minutes; and

3. to approve the final evaluation report, by applying their expert judgment to review the report of the external evaluator while keeping in mind the utility needs and views of the DIF-H consortia articulated in the DIF-H Evaluation Committee’s minutes.

An additional benefit of the External Oversight Committee was the critical review it provided to the evaluation. The feedback and technically informed commentary that the committee provided—at the planning, design, analysis, and reporting stages—improved the quality of the evaluation well beyond the independence that it engendered.

**KEY LESSONS LEARNED AND HIGHLIGHTS**

The dual committee process yielded the successful completion of the summative evaluation. In the view of the primary evaluation user, the evaluation was a useful decision-making tool not only because of the quality and utility of the evidence produced but also because of the independent perspective it provided. In the following sections, we highlight some key learning from the process.

**Independence does not imply isolation**

The conventional view of independent evaluation relies mainly on a variety of mechanisms to separate the evaluator and the evaluand. Reducing interference and capture with these approaches provides the benefit of independence, but it has the disadvantage of a loss of utility (and, arguably, accuracy) via participation. We believe the governance process used in the DIF-H evaluation is a productive
means of approaching this challenge, as it allows stakeholder participation in
evaluator guidance and real time and structured interaction between the two,
while ensuring independence via external oversight. In simple terms, we external-
ized oversight and not the evaluation.

In addition, the experience related to the external oversight committee pro-
vided new insights on the notion of neutrality. Typically, as is the case with the cur-
cent federal Policy on Results, neutrality implies safeguarding an evaluation process
from negative impacts (real or perceived) that could derive from relationships or
interests or from preconceived ideas, prejudices, or biases. While this first layer of
neutrality is critical in practical terms, evaluations also involve multiple considera-
tions framed by previous experiences, time frames, budgets, or personal expecta-
tions. These preconceived notions, while not raising any moral issues in themselves,
may sway some key decisions during the various stages of an evaluation. Since it is
firmly grounded in the realm of applied research, evaluations do require multiple
trade-offs and compromises, which must be managed in such a way as to ensure that
the end result still satisfies the information needs of its intended users. Having the
External Oversight Committee provided a more detached perspective, less framed
by these practical considerations and limitations and more focused on the overall
purpose of the evaluation. This has proven particularly beneficial since the External
Oversight Committee acted as a decision-making body, having essentially a veto on
all key stages of the evaluation. In this sense, the External Oversight Committee’s
views could not simply be considered; they had to be satisfactorily addressed. In our
view, this created a beneficial dynamic.

One notable challenge relates to the communication between the two com-
mittees. We elected to facilitate communication via written minutes rather than
direct dialogue. This approach favoured independence and, even more so, the
perception of independence, but it did not adversely affect the Evaluation Com-
mittee members’ (vested stakeholders) ability to engage productively and fairly.
Although this process was agreed to have been impartial and constructive, we
note that both Evaluation Committee members and External Oversight Commit-
tee members concurred that there might have been alternative benefits derived
from direct interaction.

Value in consultation
In our case, the regular interaction between the evaluation stakeholders and the
evaluator was found to be an active ingredient of success. The two-way flow of
information (between the consortium and the evaluator and vice-versa) allowed
the evaluator to be better informed and the evaluation results to be more targeted,
accurate, and useful. Our process adds to the body of literature and practice af-
firming this outcome. Nevertheless, it was not always easy. Formally structuring
evaluator–evaluand interactions through a committee required planning and ef-
fort. We had to be much more adaptive than planned to ensure that the timing was
right for input from Evaluation Committee members. Now concluded, we con-
tend that the four stages (planning, design, findings, and conclusion) earmarked
for interaction were appropriate and necessary, but, in practice, the committee emerged as a much more ad hoc meeting place and flexible entity than originally planned. We would encourage this too.

The unexpected benefit of consultation came from the External Oversight Committee. Although these representatives were selected for subject matter expertise, the productive insight they provided, at critical stages, was more instrumental than was expected and planned at the outset of our process. For example, at the evaluation design stage, the External Oversight Committee went beyond rejection or approval of the evaluator’s protocol (their mandate), providing significant expert insight on the framing of the evaluation study for a positive reception by our primary intended user. This input was fed back to the Evaluation Committee and the evaluator and was certainly a contributor to the successful reception of the evaluation and its eventual use. Another advantage of the committee was that all advice was from “outside the consortium box.” We learned that fresh perspective was useful not only for independence but also for innovation.

**Good governance requires resources**

The dual committee process required resources in the shape of extra time on the part of the evaluation manager (IDRC), the evaluator, and the committee members themselves. Of course, this time required financial compensation, so the process ate into the overall evaluation budget and the voluntary time committed by External Oversight Committee members. The clear position of IDRC as the “secretariat” managing the two committees likely eased the resource burden in total through the reduction of transaction costs, but it did require extraordinary attention from IDRC with respect to the evaluation. Therefore, we have noted clearly that there was a resource allocation trade-off between other evaluation activities and evaluation governance activities. In our case, this was appropriate, independence and inclusion were paramount and so we invested in them. Doing so did not interfere with the production of a quality evaluation. Others should consider this decision in their own context.

**IMPLICATIONS FOR FUTURE EVALUATION PRACTICES**

The dual committee (internal-external) governance described in this manuscript was judged to be a success by all parties involved—users, evaluators, and internal and external governors alike. In short, independence and inclusion were appropriately balanced to suit the needs of the evaluation user and also to benefit the evaluation stakeholders. This being said, we do not argue that replication in different contexts would necessarily yield the same results.

From our experience, the conditions to support successful implementation would include:

- careful construction of committee membership;
- clear communication of committee roles and responsibilities;
• transparency of committee activities;
• appropriate resourcing; and
• an enabling environment that was willing to support the process.

Under these conditions, we estimate that this model can be adapted by others seeking both independence and inclusion in an evaluation. Adaptation is emphasized since we strongly believe that contextual tailoring would be essential. The approach is not a blanket solution. Considerations of values, objectives, and the supporting conditions outlined above (at a minimum) should be thoughtfully reviewed before implementation. Nevertheless, government, international agencies, and philanthropy provide settings that we believe could benefit from the process we have outlined in this practice note.

NOTES
1 The authorship structure of this practice note is important. We each held a differing role in the management, oversight, and eventual use of the DIF-H summative evaluation. Co-authoring this practice note allowed an iterative reflection space to emerge, and we approached this note as a documentation of our after-action review. The process allowed for fair input from the various actors involved. The process was independently facilitated. Co-authorship has empowered each of us independently and our group as a whole.


4 Laffont and Tirole (1991) outline a concept of “capture” in which a public agency is coerced by a group or individual who has high stakes interest in the decisions and actions of the public agency and, therefore, invests resources into achieving control over the public entity or particular actions it takes. In economics jargon, this is known as a form of “government failure.” The field of public choice economics has built a large body of empirical and theoretical literature on this topic.

REFERENCES


International Development Research Centre. (2012). *Identifying the intended user(s) and use(s) of an evaluation (Evaluation Guideline)*. Ottawa: International Development Research Centre.


**AUTHOR INFORMATION**

Robert K. D. McLean is senior program specialist in evaluation at Canada’s IDRC; PhD student in the Faculty of Medicine at Stellenbosch University, South Africa; and non-degree track in poetry and skateboarding. He is increasingly undisciplined.

Kaitlyn Finner is the recipient of an IDRC Research Award. She is currently a policy analyst for the Nunatsiavut Government.
François Dumaine is a partner at PRA Inc., a firm specializing in program evaluation, and a past president of the Canadian Evaluation Society.

Kathryn Graham is the executive director of performance management and evaluation at Alberta Innovates, which is a Canadian-based, publicly funded provincial research and innovation organization. She has over 20 years of strategic evaluation experience in health care and health research and innovation.

Jocelyn E. Mackie is vice-president of operations and general counsel at Grand Challenges Canada, where she oversees the Finance, Operations, Legal, and Communications teams. She has been called to the Ontario Bar and the New York State Bar.

David Peckham is the chief audit and evaluation executive at the Canadian Institutes of Health Research, the Government of Canada’s health research investment agency. David has worked in evaluation for over 15 years in the public and private sector in Canada and the United Kingdom.

Kathryn Wehr, senior program officer, joined the Robert Wood Johnson Foundation in 2010. Since joining the Foundation, Wehr has worked to identify and create the conditions that allow communities and their residents to reach their best possible well-being, especially for children and their families. The Robert Wood Johnson Foundation is the largest foundation in the United States dedicated solely to improving health.

Lisa Woodward is special advisor to the President and the former director of policy and evaluation at Canada’s IDRC. An executive in the Government of Canada for 20 years, Lisa has devoted much of her career to international and sustainable development.