Valid Findings in Culturally Diverse Evaluations

Presented by Gunter Rochow

The Issue. Canadian Heritage introduces its website discussion of multiculturalism by saying that: Canadian multiculturalism is fundamental to our belief that all citizens are equal. Multiculturalism ensures that all citizens can keep their identities, can take pride in their ancestry and have a sense of belonging. While it is true that all Canadian citizens are equal in terms of their rights and obligations, and that Canada endeavours to create an inclusive society, the fact that people of different cultures can keep their identities, means that they are different. It is these differences that can have an impact on evaluations which we undertake in Canada, or abroad regarding our international political, economic, cultural or development activities. While developing this presentation, I have consulted with, and benefitted from the views of several officials of the Government of Canada.

To introduce the subject of “culturally diverse evaluations”, we must first understand what we mean by culture, and there are several meanings. Generally, we think of culture as the systems of belief, values, norms, institutions, language, historical roots by which people live. These characteristics shape how people feel, how they communicate and what they consider to be proper.

- The unique cultural characteristics of people can impact evaluations in different ways: First, in terms of communication, the evaluator must not only manage words, but be aware of the potential of messages conveyed by means of actions;
- Second, more fundamentally, there may be cultural barriers that limit trust, a situation that can restrict the effectiveness of evaluations; and,
- Third, different cultures may be used to doing things differently, which may require a change in practice to render an evaluation effective.

When we worked in Russia, we were invited to dinner in our Russian counterpart’s home. As we do in Canada, we wanted to take flowers to show our appreciation. So we stopped at a flower shop to buy a dozen roses. The shopkeeper responded with nyet, i.e. “no”. She wanted to sell us eleven. We considered that as being stingy, why not a dozen. Later we learned that in Russian culture one gives an uneven number of flowers. A bouquet of even numbers is considered to be an insult. Fortunately, our Russian host knew that we were neither insulting him nor his wife and he overlooked our faux pas.

While working in Indonesia, we noted that the Indonesian language has at least two words for ‘no”: tidak and belum. To Indonesian ears, tidak sounds harsh and confrontational. That is why many Indonesians prefer to respond with belum, which means “not yet”, leaving the conversation partner with a sense of hope that something may happen in future in spite of the present refusal. All the while the person using belum is really saying “no”, and an effective evaluator must know that this veiled refusal definitely means “no”. As a lecturer in the Harvard Workshop in Egypt 1995, I was
surprised when the course organizer came to me rather bashfully wondering whether I would mind starting my course 15 minutes later than planned and end 15 minutes later, explaining that “the people have not finished praying yet”. Are we flexible enough in our evaluation activities to modify our plans so that people can pray on their terms and at their time?

In order to reduce the potential for ineffectiveness in cross-cultural evaluations, there is a growing emphasis on the cultural competence of the evaluator. Does the evaluator understand the culture in which the evaluation occurs, and is he or she an active listener, learner and an evaluating participant within the cultural community, who demonstrates the flexibility and innovation that may be necessary to foster valid evaluation findings?

The Problem: That leads us to the question what we need to do in order to ensure that evaluations in culturally diverse environments will yield valid and reliable data, while ensuring that culturally diverse evaluations in the public sector comply with the focus on results. While important evaluation principles must not be compromised, the evaluation findings can be influenced by differing societal values, terminology and the choice of strategy. Evaluation results can be distorted unless critical cultural aspects are controlled during evaluation planning, data gathering, data analysis and reporting. Are we agreed on that?

Elements for a Solution. In recent years, the most important evaluation principle adopted in Canada, was the focus on results. To some extent, that emphasis is also found in large international institutions, such as The World Bank and the Inter-American Development Bank. Results-based approaches may be acceptable in different cultural settings when the nature of results is understood; however, its actual acceptance is often restricted by the capacity to do so. A major recommendation in Capra International’s Evaluation of the Integrated Framework [IF] for Trade-related Technical Assistance to the Least Developed Countries, in 2003, was to adopt a results-based approach. At that time, certain representatives of donor countries at the IF Steering Committee and the IF Working Group considered the results-based approach to be too burdensome for some IF participants, since the capacity may not have existed to a sufficient extent. Consequently, that recommendation was not acted upon at that time. Nevertheless, in 2004, The World Bank Operations Evaluation Department, in an independent evaluation of the World Bank’s approach to global programs, endorsed our recommendation, recognizing that a results-based approach is critical to achieving the expected impact of the Integrated Framework. It is evident that the evaluation culture focusing on results is still evolving in both developed and developing countries, as well as in multilateral institutions.

In a conventional western environment, planning for an evaluation seems to be relatively straight-forward. We identify the objectives of the evaluation, the type of evaluation, i.e.

---

http://lnweb18.worldbank.org/oed/oeddolib.nsf/24ce3bb1f94a11c85256808006a00460b4878ca5c5e3cbd85256f64005d19cf/$FILE/gppp_if_wp.pdf#page=12
formative or summative (now being renamed to “implementation” and “impact” evaluations), we design the evaluation framework, including a program description, logic model, the evaluation issues, for each issue the questions, for each question the indicators, the data sources, both secondary and primary, the timeline, the staff assignments, etc.; then we design the tools, and away we go to completing the evaluation.

In a culturally diverse environment things are not quite so simple. We must not only have a good method, but, fundamentally, as a first requirement, we must make sure that we thoroughly understand the culture or cultures involved and that we develop and maintain a good relationship with the interviewees. To achieve that, a thoroughly participatory environment for the evaluation is essential.

The evaluation literature reports an excellent example of the recognition of unique cultural values as part of the evaluation process. It is recorded in an evaluation relating to New Zealand’s criminal justice system involving Maori offenders. This experience recognized that cultural differences matter in understanding how offences may have come about and how they can be prevented in future. The methodology required a thorough understanding of Maori culture; the participation of the stakeholders in strategy design; the recruitment of researchers who thoroughly understood Maori culture and who were prepared to execute their research in a culturally sympathetic manner; the involvement of the parties in data analysis of the findings, in reporting and eventual implementation of emerging recommendations. The key criteria for successful planning of culturally sensitive evaluation must be the provision of sufficient time and capacity development within the participating cultural community. These factors also imply the need for a sufficient budget.

The planning of logistics must also take unique religious factors into consideration, as may be the case in Islamic environments. For instance, normal business activities cease about Friday noon, so that people can go to their prayer services. Similarly, western evaluators must know that during Ramadan, people fast from sunrise to sunset. Western evaluators must be aware of these cultural features, recognizing that fasting people may have special needs. Careful evaluation planning is needed, either to avoid Fridays, Ramadan and other religious holidays, or, if unavoidable, to schedule functions carefully with these religious factors in mind. The same applies to Christian and other traditions.

Evaluation planning may also need to take into consideration that in many developing countries, governmental institutions may look like Canadian institutions in terms of their structural organization, but a fully professional public service may be challenging. Many employees may be political appointees who come and go at the pleasure of the Minister. Governance issues such as political appointments, high turnover and corruption can lead to uncertain project continuity. To function in such an environment, evaluators should ensure that service contracts are signed at as high a level as possible, that work is regularly documented by means of progress reports, and that all work is signed off and approved as specified milestones are reached. To cope with such eventualities, evaluators

would do well to ensure the inclusion in their planning of capacity development and the use of multiple lines of evidence, including several key informants with different backgrounds.

The **second requirement** is to gather data in a culturally acceptable manner. In the Maori example, this took the form of letting the accused tell his or her story in a culturally safe environment, accompanied by tears and laughter. In other cultures, gender considerations may be uppermost, when scheduling interviews or participation in focus groups. In such cultures it may not be acceptable for men to interview women, or vice versa. As in the Maori example, it may be advantageous to recruit qualified evaluators who share, understand or are sympathetic to a given cultural group, if such are available, and who are capable of focused listening and observation before they speak or act. In the Maori example, the evaluation also involved engaging families in the research and carefully developing processes of communication, contact and consent. In the context of culturally diverse evaluations, data gathering efforts involve participants from groups with different priorities, histories, traditions and standards of behavior. In many instances, these cultural differences can and have contributed to miscommunication and misunderstanding, mistrust, conflict, delay, resulting in the ultimate failure of the evaluation efforts. Field evidence during the *Integrated Framework* evaluation and a *Mexican Mobility* project, show that the best way to address cultural issues is to develop effective and trustworthy working relationships among individual participants working face to face. Effective personal relationships bring about confidence that others can be counted on to fulfill their commitments. That is a level of confidence that cannot readily be achieved in any other way.

The **third requirement** is to analyse data from a cultural perspective. One reason why, in the Maori example, researchers with cultural affinity were selected is the expectation that they would be able to pick up signals that are meaningful in terms of distinct cultural values, whereas cultural outsiders might not be able to do so. In the Maori evaluation, all stakeholders were involved in the data analysis. As in all data analyses, it is important to distinguish between fact, opinion and significant cultural interpretative signals. Primary data should be *validated* to ensure comprehension before preparing draft reports. It should be noted that evaluators’ cultural identities shape their belief systems, symbolism, emotions, preferences, thinking and reasoning, all of which contribute to data analysis. Qualitative data content analysis might emphasize verbal interpretation according to a neutral coding structure.

The **fourth requirement** is to report data in a culturally acceptable way. In the context of Maori culture that meant sharing interview transcripts, essentially playing with open cards, which in traditional Canadian evaluation practice may not be acceptable.

Possible Objections. The approaches which we have sketched in this session may provoke the objections that they are too rich for most available budgets and impossible in the public sector, given the preferred *contractor selection method* of best value to the

---

5 See *Contracts Canada* ([http://contractscanada.gc.ca/en/chap10-e.htm#contractor](http://contractscanada.gc.ca/en/chap10-e.htm#contractor)). The other two contractor selection methods are: highest rated technical proposal regardless of price and “lowest bidder”.
crown as determined by a balanced rating of the technical and financial proposals. Indeed this can be a problem, but I believe there is also a solution.

In the context of the federal government, Requests for Proposal are usually written from the perspective of a traditional approach, as described above, within a limited timeframe and a budgetary ceiling. That may not be good enough to conduct culturally sensitive evaluations, for several reasons: (1) A participatory evaluation must not just involve consultation, which can result in tokenism, but it should empower the cultural community by allowing it to participate even in the evaluation design; (2) if, as we have said, evaluators must build a relationship of trust before effective evaluation results may be forth-coming, that cannot be forced into the customarily constrained timelines; and (3) while the evaluator needs to be constrained in terms of budgetary ceilings, it is unrealistic to expect the Bidder to specify detailed planned expenditures, because these will only be known when the evaluator develops with the participants from the cultural community the evaluation approach and work plan. Furthermore, the degree of flexibility in evaluations of federal government programs may vary, if the evaluation is directed to multicultural communities within Canada, or abroad. With respect to the latter, the Canadian International Development Agency (CIDA) encourages cultural sensitivity by appointing mixed teams, consisting of a Canadian and local evaluator. The Department of Foreign Affairs and International Trade (DFAIT) tries to build some flexibility up front in its statements of work for either an RFP or a call-up to a standing offer. Also, their standing offers provide the additional flexibility of amending the call-up to account for unforeseen circumstances. Nevertheless, it should be noted that participatory approaches in the design and conduct of an evaluation may not be applicable to DFAIT, since the majority of the evaluations of the policies and programs at DFAIT do not implicate an identifiable cultural community.

So, where is the solution? At the outset it should be said that it is quite possible that the evaluation budget resulting from participatory planning may be lower than traditional budgets, if actively participating members from the cultural community can be engaged as volunteers, and that is a realistic expectation, if the community believes in the sincerity of the partnership and feels that they are being empowered through their evaluation role. The flexibility that needs to be addressed is inclusiveness of evaluations, their methodologies, timelines and approaches to capture the results and impacts of programs and initiatives on diverse groups and populations in multicultural societies, such as Canada.

Also, it should be noted that evaluations in Canada, with the exception of some municipal and community-based non-profit organizations, tend to make use of mainstream tools, such as online surveys and focus groups; yet, while these are accessible to certain groups, they may represent major barriers to others, such as Aboriginal people, people with disabilities, ethno-cultural groups and recent immigrants. Furthermore, the pressure to use consistent tools to ensure the reliability of data and credibility of the findings, the limited budgets for evaluations to take into account additional costs related to outreach to marginalized individuals and groups, the interpretation services for participants with language barriers, the review of research instruments by culturally competent consultants, often result in a "one size fits all" approach. In addition, there are also ethics rules that have established requirements, such as signed written consent, which may not be
appropriate or familiar tools for people from non-European cultures. Nevertheless, while all of these issues need to be addressed, the preferred contractor selection method of best value to the crown should continue to be applied; however, once the successful candidate has been determined, the flexibility should exist, within the budgetary constraints of the proposal to fine-tune roles, fees and expenses.

Quite apart from budgetary and timeline considerations, there may also be the concern that participatory evaluations that empower people can lead to such undesirable effects as “high-jacking” evaluations by interest groups. To avoid that, sufficient control must remain with the Evaluation Project Authority. This concern also addresses the issue of independence of the evaluation function. In that regard, intellectual or behavioural independence in the execution of the function is more important that the structural independence which some evaluators desire within their organization.

**Outlook.** If we were to try some of these approaches, would the cultural pieces fit? Perhaps we can explore that further during the discussion period. I would very much like to hear from government officials, whether anything can be done in the Request for Proposal and the proposal preparation processes to allow for the kind of flexibility which culturally sensitive evaluations require.