The Canadian Journal of Program Evaluation Vol. 25 No. 3 Pages 75–84
ISSN 0834-1516

NAVIGATING EXPECTATIONS, VALUES AND CONTEXT: A CANADIAN EVALUATOR ABROAD

Linda E. Lee
Proactive Information Services Inc.
Winnipeg, Manitoba, Canada

Abstract: This article uses the story of a multi-year project conducted in four countries in east central and southeastern Europe to illustrate the challenges faced by an evaluator working in a social, political, and cultural context different from her own. Situations involving divergent expectations and unexpected events created circumstances where the evaluator had to shift roles and approaches to ensure the evaluation could remain on course. The evaluation was deemed problematic at various stages of implementation and more successful at other times. In 2010, 6 years after its completion, questions linger as to the degree the evaluation was indeed "successful."

Résumé : Par l'entremise du récit d'un projet pluriannuel réalisé dans quatre pays d'Europe de l'Est, cet article illustre les défis rencontrés par une évaluatrice travaillant dans un contexte social, politique, et culturel différent du sien. Des situations impliquant des attentes divergentes ainsi que des événements imprévus ont généré des circonstances particulières au sein desquelles l'évaluatrice a eu à changer de rôles et d'approches afin de s'assurer que l'évaluation puisse se poursuivre. À divers stades de sa mise en œuvre, l'évaluation a été jugée problématique, tandis qu'à d'autres moments, elle a été considérée un succès. En 2010, 6 ans après la fin de cette évaluation, des questions demeurent relativement à son véritable niveau de « réussite ».

INTRODUCTION

This story concerns my first evaluation experience in east central and southeastern Europe, representing a journey of self-discovery through working in a social, cultural, and political context that was clearly not my own. Since that time I have begun to understand that I am frequently working in contexts that are foreign to me

Corresponding author: Linda E. Lee, Box 439, Sandy Hook, MB, Canada R0C 2W0; <linda@proactive.mb.ca>

although, on the surface, they appear familiar. The four-year evaluation involved situations of divergent expectations, unexpected events, and complex contexts wherein I had to shift roles and approaches to ensure that the evaluation could be kept on course.

The evaluation was of a pilot project designed to address the issue of educational equity for Roma[1] children and youth in east central and southeastern Europe. Roma populations in many European countries (and indeed around the world) face both overt and systemic racism and discrimination. Impoverished living conditions, poor employment possibilities, high dependence on social assistance, and lack of access to health care and education have created situations in which Roma are an "underclass" who do not enjoy the same freedoms and state benefits as other nationals. Human rights violations are common. In the late 1990s a group of Roma parents took the Czech government to the European Court of Human Rights (*D.H. and Others v. Czech Republic*) because their children were being inappropriately assessed and placed into "special schools" for what were termed "mildly mentally handicapped" children. Countries in the region typically lack national definitions of disability (UNICEF, 2005) or use definitions in which some form of disability is connected to the sociocultural background of the child (Organisation for Economic Cooperation and Development, 1998), thus leaving the door open to discriminatory practices. The disproportionate placement of Roma children in special education is well documented, with Roma children being approximately 15 times more likely than other children to be placed in special schools (Ringold, 2001; Save the Children, 2001). In response to this phenomenon, a large American foundation decided to implement a pilot project in four countries (Bulgaria, Czech Republic, Hungary, Slovakia), based on the premise that many Roma children suffered from economic disadvantages and social discrimination rather than cognitive impairment.

The project used proven early childhood pedagogy, regular curriculum (not the simplified curriculum used in special schools), high expectations for student success, and enhanced connections with Roma families to create learning conditions supporting Roma children's school success. The goal was to integrate as many Roma children from the Bulgarian, Czech, Hungarian, and Slovakian pilot classes into mainstream education by Grade 4, following three years of project implementation, with the intervention beginning when the children were in Grade 1. In fact, the project and its evaluation were extended to include a fourth year when the majority of children were integrated into mainstream Grade 4 classes.

As the evaluator on the international implementation team, I worked in consultation with NGO directors and researchers from the region. My first task was to create the evaluation framework. Then I was charged with managing the overarching evaluation, while researchers in the country conducted activities necessary to satisfy governmental requirements in their particular countries. My involvement came about because of a chance meeting with the woman who was the project director. Two mutual friends recommended me as someone who had knowledge of both program evaluation and student assessment. I agreed to come to the initial meetings in Budapest to learn more about the project and to develop a multifaceted evaluation framework that would monitor implementation and assess project outcomes. Then both the project director and I would decide if we wanted to continue to work together for the duration of the project.

This article tells the story of my attempts to use a participatory approach both in developing the initial framework and in coordinating the multicountry evaluation over a period of four years. At times it was successful, but at other times evaluation-related issues jeopardized the project itself. Ultimately, the success of this evaluation can be questioned in light of the fact that very little has changed for Roma children and youth in the region, despite the evidence provided by the evaluation which was used in the European Court of Human Rights and widely disseminated.

THE EVALUATION EXPERIENCE

My evaluation experience began on a hopeful note. I was included in the first meeting of the teams who were implementing the project in each country. The evaluation and the project were going to start at the same time! Each country team (spearheaded by an NGO working in the area of early childhood education) included the NGO director, a master teacher trainer employed by the NGO (who would work with the teachers in implementing schools), and a recognized researcher chosen by the NGO from within their own country. I was part of the international coordinating team whose composition paralleled the country teams with the project director, a trainer, and a researcher/evaluator. In addition, our international team included (at least at the outset) a Roma academic and activist. All participants were to speak either English or Russian in order to facilitate communication.

I arrived at this first meeting with the understanding that the intent and parameters for the project and the evaluation would be discussed. I was unprepared to learn during the first morning that I would have the researchers for the afternoon session. Consequently, I spent the lunch break sitting on a bench under a rose arbour in the park surrounding the hotel trying to stay calm and figure out what I was going to do with the researchers. I decided to proceed as I would in North America, using a collaborative approach, engaging the researchers in a process of developing a foundation for the evaluation framework. Although the discussion had its quiet moments, I attributed this to the fact that I was facilitating through an interpreter for the first time. I concluded our session by outlining the next steps in the process, which would begin with me sending them the framework for their review and input. I did not understand that the expectation of the researchers and the NGO directors was that the international evaluator was the "expert" who had been selected because she "knew best" and would be directing the work of the other researchers. I learned later, through the international project director, that my consultative approach was viewed as exhibiting inexperience, lack of knowledge, and an inability to take charge.

While an overarching evaluation with a common methodology was the intention, each country team had to meet its own governmental requirements to be approved for a pilot project. Consequently, the researchers had to add certain elements in order to meet governmental expectations. In one case, the researchers planned to administer a battery of assessments on students entering Grade 1 that were developmentally inappropriate in their volume as well as their cultural biases. At this juncture, with the support of the project director, I did switch into a more directive role. When the researchers explained their government required them to use some nationally recognized assessments, we negotiated and finally required them to reduce the number of assessments and to exclude particular elements.

One of the most critical moments for the evaluation (and indeed the whole project) arose through an evaluation discussion. Once the project had begun, the teams were brought together for semi-annual meetings to update progress, share experiences, solve problems, and later to participate in training sessions. The meetings included times where everyone was together, as well as sessions when the NGO directors, teacher trainers, and researchers/evaluators met separately. At the meeting that occurred during the first autumn of implementation, I facilitated the evaluation session. One of the issues we needed to discuss was how we were going to identify the Roma children in the special schools and classes, as there were a few students from the majority populations in these classes, invariably children who did have some cognitive disability. We were still operating in English and Russian (although as the project progressed we moved to multiple languages for our meetings). The Roma member of our international team, who spoke English but also had a working knowledge of Russian, was participating in our meeting. We were moving toward agreement that the child's family would self-declare their ethnicity, in many cases through liaison with the Roma teacher assistant/family coordinator who was working in the child's classroom.

Then one of the researchers, who had been relatively quiet, began to speak. My Roma colleague, who could understand Russian, raised his voice in response. Their conversation was just beginning to be translated into English, so I was listening to the exchange on a delay. The researcher had explained that it was easy to tell who the Roma children were because of their appearance, including their dark skin colour and dirty clothing. My Roma colleague had responded heatedly. At this point the discussion rapidly moved to accusations of racism, vehement denials, everyone speaking at once. One of the other researchers started to cry, saying "It is not just gypsies who were victims of the Holocaust." Not only was hope of a successful evaluation evaporating before my eyes, the entire project appeared to be in jeopardy. As quickly as I could, I exerted control as chair of the meeting. I cannot recall exactly what I said, but I attempted to bring the discussion back to the central focus and point out the error of the researcher's position. Then I suggested we take a break. Going immediately to the project director, I enlisted her assistance with a group debriefing where we reviewed project expectations and established ground rules for our behaviour.

That evening the researcher apologized to my colleague, assuring him that he had learned something. He would now tell his wife not to threaten their children with being stolen by gypsies if they were behaving badly. Unfortunately, we continued to hear racist comments from this researcher, although they were less public. The coordinating team realized that anti-bias training (later known as "education for social justice") was required, not only for the educators implementing the project, but also for everyone participating on the country teams and the international coordinating committee.

Upon submitting the Year 1 report to the client (funder), the evaluation was again in jeopardy. The client did not believe the teachers' assessments of the children's progress—or lack thereof. Results from the first year also indicated that the teachers were not implementing the project as intended. Many teachers were continuing to use special education curriculum, as their expectations for children's achievement remained low. Few knew how to work with the Roma teaching assistant/family coordinator to enhance student learning, bring culture into the classroom, and connect with families and community. While this suggested to me that the teacher assessments of children's skill levels were likely accurate (and the master teacher trainers who were observing in the classrooms concurred), the client remained unconvinced. S/he also argued that, to be accepted by governments, stronger achievement data were required.

The need for more credible assessments was obvious. Consequently, the trainer on the international team and I worked with representatives of governments, pedagogical institutions, and other experts in each country to support the development of centrally developed curriculum-based assessments. In this way we had a more "objective" measure as to whether the students were meeting the requirements of the regular curriculum. Fortunately, the development of these assessments helped teachers understand the curriculum requirements, and, with intensive support from the international trainer and the master teacher trainers, implementation gains were made. By the end of Grade 2, approximately two thirds of the children were meeting the expectations of the regular curriculum in the official language and mathematics. The project and the evaluation were both viewed as successful at this juncture.

Success brought new issues. If children continued to exhibit these levels of achievement, many would be candidates for integration into regular schools in Grade 4. The project director and funder recognized that much groundwork needed to be done if the children were to be successfully integrated. Therefore, a major conference was held to which elected government officials, policy makers, representatives of Roma organizations, and others were invited to learn about the project and its results. Although the evaluation results were now being touted, I had concerns that the children's gains might not hold through to the end of Grade 3. However, if the appropriate conditions to support children's integration were to be fostered, waiting until the end of Grade 3 would be too late. The second-year results were highlighted in the conference. I breathed a sigh of relief a year later when the Year 3 results exhibited the same trends.

EXPLANATORY FACTORS

My view of evaluation includes many elements that would be situated on the "use" branch of the evaluation theory tree (Alkin & Christie, 2004) and is grounded in a participatory approach (Cousins & Earl, 1992; King, 1998; Lee & Cousins, 1995; Patton, 1997). For a participatory approach to succeed, King (1998) argues that there must be an accepting power structure; shared meaning of experiences among participants, volunteers, and leaders; enough time; enough resources; and a great degree of interpersonal and organizational trust. In this case, the official power structure included the funder and project director, while the directors of the NGOs also played an influential role. Although the project director throughout the process supported and believed in the evaluation, the funder and the NGO directors, at different points, were not accepting of the evaluation process and/or results. At the outset, the NGO directors were reluctant to believe the evaluation would be successful, that is, would provide the results they needed to gain political support in the longer term. Then, after the first year, the funder did not view the evaluation as successful as s/he did not find the results to be credible.

Lack of a shared meaning among participants threatened the evaluation, as the researchers came with different perspectives on evaluation, a situation that was exacerbated by multiple languages and divergent belief systems. The importance of language has been recognized as a key to trust building because of the need to unlock meanings that are inaccessible if underlying assumptions are not accurately shared (Hall & Hood, 2005; Mertens, 2009). Our research group struggled with language, and, with so much divergence in our underlying assumptions, interpersonal trust among the researchers was, at best, slow to build. Fortunately, at another level, the project director and I shared language and values. Trust was high between us, which lent support to the evaluation as it developed and evolved.

Project results did have short-term impact on educational policy in a number of countries, such as supporting the provision of Roma teaching assistants/family coordinators in schools with significant Roma populations. The data were used as evidence in the European Court of Human Rights in the case *D.H. and Others v. Czech Republic*. Although the case was lost, the decision was later overturned on appeal. The findings from the project also helped to shape a multi-year comprehensive educational strategy for Roma children and youth in seven countries. However, at a conference in Tirana, Albania

(March 2010) that focused on educational policy and educational equity, I was distressed to learn that very little had changed for Roma children's education. I was also disheartened that, even though the conference was sponsored by organizations related to the original funder and attended by many people connected into this network, our evaluation results were not widely known or referenced. This lack of long-term use may suggest that, ultimately, the evaluation was not "successful." In part, staffing changes within the funding organization may have contributed to reduced dissemination of results. Regarding longer-term policy impact, the potential for EU accession encouraged political attention to Roma issues, a pressure that no longer exists now that the four countries are EU members. Finally the influence of deep-seated racism (both personal and systemic) cannot be underestimated.

LESSONS LEARNED AND CONSIDERATIONS FOR PRACTICE

The lessons learned included the need to pay close attention to understanding the values of those involved, as these are crucial to understanding the context in which the evaluation will operate. Although, as evaluators, we recognize the importance of context, digging down to personal values could be a helpful factor in reducing the likelihood that an evaluation will be sabotaged or derailed. One possible strategy is to use an exercise with stakeholders early in the evaluation process that prompts discussion of values.

As we pay more attention to cultures when we cross borders, as evaluators we also need to recognize different cultures within our own, seemingly more familiar, context. Culture itself is a dense and multilayered phenomenon, the effects of which add another dimension of unpredictability to evaluation and, ultimately, the influence or success of the evaluation process and its results. The CES's Competencies for Canadian Evaluation Practice includes "attending to issues of diversity and culture," which, I would argue, should be a daily consideration for evaluators.

Recognizing that your own definition and approach to evaluation may not be shared by others (regardless of the social and political context) is key to dialogue and building the trust relationships necessary for developing and implementing a successful evaluation. Being explicit at the outset about one's approach to evaluation and describing what this approach will mean for the evaluation and for the roles of various participants can be an effective practice.

The other element that requires specificity concerns the expectations that clients/funders and other stakeholders will place upon the evaluator. What is the role or roles the evaluator is expected to play? As an evaluator, do you understand the expectations and how you can resolve the issue of divergent expectations—including your own? Reconciling these expectations again relates to the process of building community and trust among stakeholders.

Despite this evaluation having fought through credibility issues and conflicting expectations, the evaluation did experience short-term use. However, the lack of longer-term impact is troubling. Given the recent revelation that long-term use is not in evidence, a next step is being contemplated. We will be putting forward a proposal regarding a follow-up of the Roma children who were involved in the project. These children are now 16 years old. Are they still in school? What is happening in their lives? Does it appear that their later life experiences were influenced by participation in the project? If such a follow-up can be orchestrated, perhaps the impact of the original evaluation will be revitalized and strengthened with new learning.

NOTE

1 Roma have often been referred to as "gypsies," a derogatory term associated with many negative stereotypes.

REFERENCES

Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 12–65). Thousand Oaks, CA: Sage.

Cousins, J. B., & Earl, L. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*(4), 397–418.

Hall, M., & Hood, D. W. (2005). Persuasive language, responsive design: A framework for interculturally responsive evaluation. In S. Hood, R. Hopsonh, & H. Frierson (Eds.), *The role of culture and cultural context* (pp. 39–58). Greenwich, CT: Information Age.

King, J. A. (1998, Winter). Making sense of participatory evaluation on practice. *New Directions in Evaluation, 80,* 57–67.

Lee, L. E., & Cousins, J. B. (1995). Participation in evaluation of funded school improvement: Effects and supporting conditions. In J. B. Cousins & L. Earl (Eds.) *Participatory evaluation in education: Studies in evaluation use and organizational learning* (pp. 72–85). London, UK: Falmer.

Mertens, D. M. (2009). *Transformative research and evaluation.* New York, NY: Guilford.

Organisation for Economic Cooperation and Development. (1998). *Co-ordinating services for children and youth at risk: A world view*. Paris, France: Centre for Educational Research and Innovation.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.

Ringold, D. (2001). Education of the Roma in Central and Eastern Europe: Trends and challenges. In C. McDonald, J. Kovács, & C. Fényes (Eds.), *The Roma education resource book* (pp. 17–42). Budapest, Hungary: Open Society Institute.

Save the Children. (2001). *Denied a future: The right to education of Roma / Gypsy and Traveller childre*n. London, UK: Author.

UNICEF, Innocenti Research Centre. (2005). *Innocenti insight: Children and disability in transition in CEE / CIS and Baltic states.* Florence, Italy: Author.

Linda E. Lee is a partner and Vice-President of Proactive Information Services Inc., a social research and program evaluation company, based in Manitoba. Linda has been a teacher, facilitator, researcher, and evaluator, working in the evaluation realm for over 30 years and with Proactive since 1989. She has worked across Canada and in many countries in east central and southeastern Europe, as well as in Lithuania and Argentina. A former National President of the Canadian Evaluation Society, Linda received the CES Service Award in 2003 and, in 2009, was inducted as a CES Fellow. Linda is a passionate advocate for evaluation and its role in helping to make the world a better place.