Evaluating System Change Initiatives: Advancing the Need for Adapting Evaluation Practices

Cheryl Poth
Centre for Applied Measurement and Evaluation, Department of Educational Psychology, Faculty of Education, University of Alberta

Jacqueline Pei and Erin Atkinson
Department of Educational Psychology, Faculty of Education, University of Alberta

Tara Hanson
Alberta Centre for Child, Family, and Community Research

Abstract: This article describes key challenges experienced and addressed during the evaluation of the Government of Alberta’s 10-Year Fetal Alcohol Spectrum Disorder (FASD) Strategic Plan (2007–2017). The purpose was to understand how the strategic plan was progressing toward outcomes at the fifth year of its implementation. Following a description of the system change initiative and evaluation context, an account of key challenges for one outcome is presented, including attempts to address and the effects on the evaluation. The implications for evaluation practice focused on encouraging evaluation participation and enhancing usefulness of data highlight the need for infrastructure to support evaluation of system change initiatives.

Keywords: complex evaluations, outcomes evaluation, systems-level evaluation


Corresponding Author: Cheryl Poth, Centre for Applied Measurement and Evaluation, Department of Educational Psychology, Faculty of Education, University of Alberta; email: cpoth@ualberta.ca

© 2016 Canadian Journal of Program Evaluation / La Revue canadienne d’évaluation de programme
31.2 (Fall / automne), 242–252 doi: 10.3138/cjpe.263
Mots clés : évaluations complexe, évaluation des résultats, évaluation au niveau des systèmes

INTRODUCTION

Evaluators have responded to increasing interest in assessing large-scale social change with guiding practices for how to approach evaluating system change interventions (Hargreaves, 2010). Yet what remains to be further explored are the nature of challenges experienced and, perhaps more importantly, how evaluators can effectively respond to these challenges during an ongoing evaluation. To begin to address the dearth of illustrative examples, we offer the following practice note, with the purpose of advancing lessons learned for informing evaluation practice of system change initiatives. As is our usual practice as university-based researchers and evaluators, we sought and received ethical approval for studying the evaluation process in addition to the evaluation outcomes before undertaking this work.

The authors were all closely involved in the evaluation of a system change initiative in different roles aligned with individual experience and expertise that we believe are important for telling the story of this evaluation process and ultimately contributing to the evaluation practice literature. The organization tasked with implementing the mandated evaluation is the Alberta Centre for Child, Family, and Community Research (the Centre), a not-for-profit charitable corporation established in 2003 as a partnership between the Alberta government, universities, and the community, whose mandate involves promoting capacity building in research and evaluation. For these reasons confidentiality has been waived and we identify our roles. As the Centre's Director of Knowledge and Partnership Development, Hanson managed the development and implementation of the mid-point evaluation of the 10-year strategic plan. In this role, she worked with an Evaluation Advisory Committee to engage researchers and evaluators to undertake segments of the evaluation. One evaluation group was the Alberta Clinical and Community-based Evaluation Team (ACCERT) at the University of Alberta, composed of faculty and graduate students from the Department of Educational Psychology within the Faculty of Education. As the cofounders of ACCERT, Pei and Poth conduct and mentor community-involved program evaluation and applied social research with a focus on building capacity both within the University-based team and with clients and stakeholder organizations. Pei—a registered psychologist with expertise within the field of Fetal Alcohol Spectrum Disorder (FASD)—served as the primary contact with the evaluation project manager, whereas Poth—an experienced program evaluator with expertise in mixed methods, qualitative, and quantitative approaches—served as the lead methodologist. Finally, Atkinson—a doctoral student with expertise in program evaluation and FASD research—served as member of the ACCERT team of 10 graduate students.

This practice note is organized in three sections. First we provide background information by describing the system change initiative and evaluation context,
then we present the issues we encountered by describing the challenges including the resulting impediments to the evaluation process and the adaptations we made to our evaluation practice. Finally, we convey the lessons learned by describing recommendations for evaluation practice and results of the evaluation of the system change initiative.

DESCRIPTION OF SYSTEM CHANGE INITIATIVE AND EVALUATION CONTEXT

Why was the evaluation conducted? What did the client want to learn?

The Government of Alberta’s FASD 10-Year Strategic Plan (2008), intended to support large-scale social change, is the first provincial plan of this nature to be evaluated in Canada. Its unique development process involved nine partnering ministries participating on a Cross-Ministry Committee (CMC) cochaired by Alberta Health and Alberta Human Services. The plan articulates a vision for Alberta to have a “comprehensive and coordinated response to Fetal Alcohol Spectrum Disorder across the lifespan and a continuum of services that is respectful of individual, family, culture, and community diversity” (Government of Alberta, 2008, p. 7).

Strategies were developed to achieve outcomes in three service areas: awareness and prevention, assessment and diagnosis, and supports for individuals and caregivers. Specific activities were also aligned with outcomes in the areas of research and evaluation, strategic planning, education and training, and stakeholder engagement. Evaluations were mandated for Years 5, 7, and 10 of the plan by the Government of Alberta’s Standing Policy Committee on Health and Community Living to monitor targets and measure progress toward the nine outcomes established in the plan:

1.  
   a) Albertans understand that alcohol use during pregnancy can lead to FASD, that FASD can be prevented, and that FASD prevention is a shared responsibility.
   b) Alcohol use during pregnancy is eliminated.

2.  
   Adults, children, and youth suspected as being affected by FASD have access to timely and affordable diagnostic and assessment services.

3.  
   Individuals affected by FASD and their caregivers have coordinated access to support services to meet their needs.

4.  
   Service providers and families/caregivers have knowledge of and access to training and educational resources that are based on research and leading practices.

5.  
   The planning and delivery of provincial government programs and services associated with FASD are accomplished through a collaborative approach.

6.  
   Basic and applied research and findings, including those from monitoring and evaluation systems, are used to inform FASD strategic planning, FASD prevention activities, and FASD-related programming.
7. Mechanisms are in place to facilitate and encourage stakeholder engagement in the FASD-CMC strategic planning process, as well as to provide stakeholder opportunities for networking and information sharing.
8. Secondary disabilities associated with FASD and their impact on Albertans are reduced.
9. The cost of FASD to Albertans is reduced.

The Ministry of Health provided a grant to the Centre to undertake the development and implementation of the mid-point system change initiative evaluation of the 10-year strategic plan. Thus, the purpose of the Year-5 evaluation was to provide baseline data and key recommendations to assist in continuous quality improvement of the initiatives conducted under the plan and to inform subsequent evaluations to be conducted in Years 7 and 10 of implementation.

**What resources were available for conducting the evaluation, and how were they allocated?**

Funding for the mid-point system change initiative evaluation represented approximately 3% of the annual budget allocated for operationalizing the strategic plan. The Centre’s Director of Knowledge and Partnership Development convened an Evaluation Advisory Committee of key stakeholders to provide advice, enhance communication among partners, and ensure a collaborative approach throughout the term of the evaluation. To measure progress toward each outcome, it became clear that multiple distinct projects with different methodologies and expertise would be required. The project manager recruited organizations and individuals with the specialization necessary to successfully conduct and complete these projects. Twelve contractual agreements were implemented and managed within the 18-month term of the grant.

Approximately 35% of the total evaluation budget was allocated to ACCERT to answer evaluation questions related to outcomes 1b, 2, 5, 6, and 7 (see section above). The description of the challenges encountered is focused on our work related to outcome 1b because it is representative of our experiences across outcomes.

**DESCRIPTION OF CHALLENGES AND RESULTING EFFECTS AND RESPONSES**

**What challenges did you face in conducting this evaluation?**

**How did the challenges impede the evaluation? How were the challenges addressed?**

Two major data access challenges were experienced during our work on outcome 1b (see Table 1). In hindsight we realized that the challenges were applicable across most of the outcomes and largely attributable to a focus on developing infrastructure to deliver services during the first five years of the plan. This meant few resources for
data collection and management. The resulting absence of standardized data collection processes and lack of database development were identified as the first significant challenge for the evaluation. Furthermore, Evaluation Advisory Committee members agreed that some of the plan’s intended outcomes were not measureable because indicators had not been defined and baseline data had not been generated; this is why providing baseline data became a key purpose for the evaluation. The second significant challenge for the evaluation emerged as a result of the efforts

Table 1. A Summary of System Change Initiative Evaluation Challenges for Outcome 1b

<table>
<thead>
<tr>
<th>Description of challenge</th>
<th>Encouraging Evaluation Participation</th>
<th>Enhancing Usefulness of Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Program service providers reluctant to participate in the evaluation, as evidenced by low response rates to initial data collection.</td>
<td>• Outcomes were not measureable as they were initially defined.</td>
<td></td>
</tr>
<tr>
<td>• Lack of data coordination and management.</td>
<td>• Lack of data coordination and management.</td>
<td></td>
</tr>
<tr>
<td>• Consulted with service providers about nature of concerns.</td>
<td>• Worked with an Evaluation Advisory Committee to explore and agree on proxy measures and sources of data.</td>
<td></td>
</tr>
<tr>
<td>• Partnered in developing data procedures that were relevant to their work, respectful of their time, and enabled timely use of findings.</td>
<td>• Focused on generating findings that were meaningful for participants and could be used for improvement.</td>
<td></td>
</tr>
<tr>
<td>• Implemented innovative data collection strategy using appropriate methods for participants.</td>
<td>• Engaged in the collaborative development of recommendations.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Effects on evaluation engagement:</th>
<th>Catalyst for evaluation use:</th>
<th>Facilitator for evaluation use:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Increased participation using appropriate methods to create validating experiences.</td>
<td>• FASD-CMC subsequently developed an action plan to respond to Year 5 evaluation recommendations with infrastructure for subsequent Year 7 and 10 evaluations.</td>
<td></td>
</tr>
<tr>
<td>• Greater interest attributed to generating and communicating rich understandings of role in collaboration and prevention.</td>
<td>• An outcome-based management plan, which was adopted by the FASD-CMC as its Strategic and Operational Plan.</td>
<td></td>
</tr>
</tbody>
</table>
of multiple individuals and teams collecting data from participants within a short time, which resulted in several cases of participant fatigue. The following section describes what happened when we attempted to generate data specifically to address outcome 1b (Alcohol use during pregnancy is eliminated), including the need to use proxy measures and respond to participant concerns.

The Evaluation Advisory Committee defined the evaluation question for outcome 1b: “Is there evidence that alcohol use during pregnancy has been reduced or eliminated among at-risk women and overall among women in Alberta?” At-risk women were identified as those who have participated or are currently participating in the Parent Child Assistance Programs (PCAP) in Alberta. The program served as a proxy measure for outcome 1b. The program was designed to prevent drug and/or alcohol-exposed live births among high-risk mothers (Grant, Streissguth, & Ernst, 2002) who have substance abuse issues and are disconnected from community services; these women are exceptionally vulnerable and are in dire need of support and resources (Ernst, Grant, Streissguth, & Sampson, 1999). By focusing on reducing risk and emphasizing the client-advocate relationship, PCAP is distinguished from other interventions that force women to remain completely abstinent from drugs and/or alcohol (Ernst et al., 1999).

The Advisory Committee was specifically interested in generating an in-depth understanding of the programmatic experiences of the PCAP service providers who work with women engaged in high-risk behaviour to prevent future births of children prenatally exposed to drugs and/or alcohol. Staff reluctance to participate became apparent when a review of the data collected to the mid-point evaluation revealed a very low response rate. Early attempts to engage participants involved consultations and becoming aware that their diminished interest had resulted from their recent evaluation experiences. For example, program service providers voiced frustration that despite their having completed several surveys during the previous two years, findings were never shared with them. Moreover, there were complaints that the previous surveys had been lengthy, repetitive, and perceived as having no relevance to their PCAP work.

To begin, we partnered with a PCAP manager to discuss ways in which program service providers might participate in more appropriate data collection efforts as key stakeholders. Priority was given to gathering information that would be of value to participants, provide immediate access to findings, and limit the time necessary for participation. Together, an innovative exercise was developed as part of a PCAP organizational retreat involving the integration of a visual data collection method (i.e., quilting) with a more traditional oral data collection method (i.e., focus groups). In adherence to principles guiding research ethics, although all program service providers were invited to participate in a quilt-making exercise and focus group discussion, only those who gave explicit consent were included in the data analysis (for full description of procedures, see Job et al., 2014). A facilitator conscientiously guided the quilting process with a question intended to be individually interpreted, saying, “Using words or illustrations, tell us about your experiences in PCAP.” In response, participants created individual
quilt squares (later compiled into a large quilt) and, in some cases, added words to complement their visual expression. In recognizing that, for participants, this exercise was as much about the process as the product, photographs were also taken during this task.

The atmosphere of the focus group was similarly carefully constructed to support sharing of experiences with the aim of networking between regions, using such guiding questions as “Tell us your story about how the PCAP program has changed over the past five years.” The data collection approach had a catalyzing effect resulting in three important impacts on the evaluation. First, in addition to high rates of participation, the vast majority of the 48 participants (more than 95%) agreed to have their data included in the evaluation. Participants attributed their willingness to participate because the data collection methods had helped them to convey their experience in a relaxing and creative environment. At the end of the day, several participants described the focus groups as being validating because they revealed shared and similar experiences among other program service providers across regions. In particular, participants reported that this helped them to realize that they were members of a community and therefore “not alone.”

Second, the integrated findings generated a rich understanding of the program experiences related to collaboration and prevention from the perspective of the PCAP service providers that may not have been otherwise accessible. Specifically, whereas the quilting enabled each participant to share his or her story through conversation and creative expression, the focus group offered perspectives across groups of service providers that were useful for their own interest and for the evaluators to gain in-depth understandings of human connections (Ansay, Perkins, & Nelson, 2004).

Enhancing data usefulness by addressing stakeholder interest in the evaluation process and findings created opportunities for stakeholders to build personal connections with the data—known as the personal factor—and continues to be highlighted in use-focused approaches adopted by the client (e.g., Patton, 2011). In keeping with an appropriate approach, findings were shared across mediums and tailored to audiences. For example, formal presentations were made to government officials, whereas a video presentation was created for participants. In addition, the completed quilt was given to the PCAP organization, and a plaque with a photo of the quilt was given to each region. The anticipated long-term outcomes are enhanced relationships with program service providers and increased use of the evaluation results.

**RECOMMENDATIONS AND RESULTS**

An emphasis on data usefulness and stakeholder engagement throughout the evaluation process was critical to the success of the project. The initial lack of data available and participant willingness presented challenges to effectively measuring specific outcomes; this required collaboration to agree upon acceptable proxy measures and relevant sources of information. Early engagement of stakeholders
is also recommended in the planning stage in addition to developing recommendations based on the evaluation results to address policies, practices, and continuous improvement of the FASD service network model. FASD-CMC subsequently developed an action plan to respond to Year 5 Evaluation recommendations, which included providing clients with assessment for intervention and wrap-around services, defining sustainability, clarifying outcomes, developing a data-collection model, further developing governance structures, improving the funding model, and increasing access to programs. Following the evaluation, FASD-CMC worked with the networks and consultants who participated in the evaluation to articulate clear, measurable outcomes for both clients and the system. These adaptations to evaluation practices were brought together in the form of an outcome-based management plan accompanied by the introduction of online reporting systems. While the strategic direction and targets identified in the FASD 10-Year Strategic Plan did not change, how success is measured and evaluated evolved significantly.

What, if any, are the big-picture issues that the evaluation community should address?

Challenges related to coordination and collaboration may be expected during outcome evaluations of system change initiatives, yet our experiences appear to be intensified by the initial reluctance of program service providers to participate and by the lack of initial focus on the data collection experiences of participants. Specifically, it was the project manager’s adept listening and response to concerns that was critical to the ultimate success in creating an evaluation process and generating findings that could be used for both improvement and increased accountability by stakeholders. We believe further examples of effective practices and successful strategies from the evaluation community will better position evaluators to respond to challenges inherent in evaluations of systems-level outcomes. To begin, we propose the following actions:

- Articulate how evaluative procedures can be translated from strategic-plan-level aspirational outcomes to program-level data collection initiatives. Operationalizing outcomes is crucial for informing delivery of services across diverse fields from homelessness and child custody cases to education and business (e.g., Austen & Pauly, 2012; Kelly & Ramsey, 2009). What is missing from the literature are guiding examples of evaluation questions that use diverse data-collection procedures that are appropriate for stakeholder groups. Translation skills are essential for ensuring that data collected within programs are representative of stakeholder interests and contribute to the client’s informational needs. Indeed, attention to these skills is warranted within the Canadian context as the focus of one of the Competencies for Canadian Evaluation Practice within the Situational practice domain; specifically, Competency 3.4 points to the need for identifying the interests of all
stakeholders (see Canadian Evaluation Society, 2010). A specific focus on identifying and prioritizing the interests of key stakeholders (i.e., clients) while balancing the interests of other stakeholders (i.e., program service providers) involved in the evaluation were vital to the current evaluation.

- **Describe how coordinated data-collection and analysis strategies can be developed that are respectful and responsive to stakeholder groups.** The dependency of accurate inferences on generating data from reliable and trustworthy sources is well established (e.g., Creswell, 2012). Missing are examples highlighting how to involve stakeholders in the process of developing and implementing data-collection strategies in ways that address specific issues and ultimately increase the potential for use and interest in the evaluation. Increased use and interest in the evaluation findings and process from participating in the evaluation activities is well established in the literature as the personal factor (Patton, 2008; 2012). Indeed, attention to engaging stakeholders in decisions is warranted within the Canadian context as the focus of one of the Competencies for Canadian Evaluation Practice within the Technical Practice domain; specifically, Competency 2.8 highlights the need for identifying data sources (see Canadian Evaluation Society, 2010). A specific focus on consulting with stakeholders to identify appropriate data sources and how to best address issues in accessing these sources were essential to the current evaluation.

- **Explore how evaluation capacity can be catalyzed for meeting the informational needs of diverse audiences.** Building capacity for integrating interpretations of programmatic data that can be used for assessing system-level outcomes is important. What is missing in the current efforts to define the dimensions of evaluation capacity-building is consideration for collaboration among individual evaluators (e.g., Bourgeois & Cousins, 2013; Preskill & Boyle, 2008). Indeed, attention to these skills is needed within the Canadian context as the focus of one of the Competencies for Canadian Evaluation Practice within the Interpersonal Practice domain; specifically, Competency 5.8 calls for collaborating and partnering skills (see Canadian Evaluation Society, 2010). A specific focus on coordinating and working toward mutual benefit among partners were key to the current evaluation.

**REFERENCES**


Evaluating System Change Initiatives


**AUTHOR INFORMATION**

Cheryl Poth, PhD, is an associate professor in the Centre for Applied Measurement and Evaluation within the Department of Educational Psychology, in the Faculty of Education at the University of Alberta and a cofounder of the Alberta Clinical and Community-based Evaluation Research Team (ACCERT). She teaches and conducts research on program evaluation, working with stakeholders from school boards, postsecondary institutions, and federal organizations in the areas of educational programs and health services. She is a regular contributor to the Canadian and American Evaluation Association conferences.

doI: 10.3138/cjpe.263
and publications. Her research interests include evaluation use with particular emphasis on developmental evaluation and evaluation use.

**Jacqueline Pei**, RPsych, PhD, is an associate professor within the Department of Educational Psychology, in the Faculty of Education at the University of Alberta and a cofounder of the Alberta Clinical and Community-based Evaluation Research Team (ACCERT). A practicing Registered Psychologist for the past 12 years, Dr. Pei began her career as a criminologist and forensic counsellor working with incarcerated youth. Motivated by this early work, she returned to academia to study youth at risk, child development, and neuropsychology, leading to her current focus on interventions for individuals with Fetal Alcohol Spectrum Disorders. Building on her work with various community and government agencies, including serving on the FASD Clinical Diagnostic Team at the Glenrose Rehabilitation Hospital, Dr. Pei currently leads the Intervention Network Action Team (iNAT) for the Canada FASD Research Network and is cochair of the Services and Supports Council of Alberta, roles that facilitate the link between research, policy, and practice.

**Erin Atkinson**, MEd, is a doctoral student in the School and Clinical Child Psychology (SCCP) program in the Department of Educational Psychology, in the Faculty of Education at the University of Alberta. She developed an interest in program evaluation after taking a course with Dr. Cheryl Poth, and has since been involved in designing and implementing several program evaluations with the Alberta Clinical and Community-based Evaluation and Research Team. She has experience in both quantitative and qualitative data collection and analysis, and has a specific interest in mixed methods research design. Erin is interested in the area of FASD prevention and the education of children with FASD. Her research focuses on preparing preservice teachers to feel confident and efficacious in working with students with FASD in their future classrooms.

**Tara Hanson**, MACT, is Executive Director of Business and Partnership Strategies with the Alberta Centre for Child, Family, and Community Research. She enjoys working across sectors with academics, policy makers, service providers, and community agencies. Her areas of expertise include system-level evaluation, project management, stakeholder engagement, and knowledge mobilization. In addition to managing the Year 5 and Year 7 evaluations of the Government of Alberta’s FASD 10 Year Strategic Plan, she recently led the development of the Housing and Homelessness Research Strategy for Alberta, and the Child Intervention Service Quality Framework. She is currently supporting the Centre’s Aboriginal Advisory Committee to develop an Aboriginal Research Strategy. Tara has a BA and Master of Arts in Communications and Technology from the University of Alberta and a Knowledge Translation Professional Certificate from the University of Toronto, Faculty of Medicine. She recently completed the Innovation Leadership Program through the Waterloo Institute for Social Innovation and Resilience and the Haida Gwaii Higher Education Society.