

Utilization Focused Developmental Evaluation: Learning Through Practice



ISSN 1556-8180
<http://www.jmde.com>

Ricardo Ramírez

Independent researcher and consultant based in Guelph, Ontario, Canada

Galin Kora

Independent researcher and consultant based in Ottawa, Canada

Daniel Shephard

Research and Curriculum Manager with the Aflatoun Secretariat, Amsterdam, The Netherlands (the client)

Background: Utilization-focused evaluation provides an overall decision-making framework with the intention of ensuring evaluation products and processes are actually used. Developmental evaluation provides a structure to learn from an experiment or pilot in the making and provide feedback to course-correct and improve the ongoing effort. In this paper we report on a project where we combined both into a utilization-focused developmental evaluation (UFDE).

Purpose: To determine the theoretical complementarities and the practical advantages of combining UFE with DE, by reflecting on a practical experience. We include a synopsis of the methodology along with a sample of findings, followed by a reflection of the overall process. We emphasize the conditions that enabled this experience to evolve to guide other practitioners interested in this learning approach to evaluation.

Setting: The context was the piloting of a social and financial education curriculum for youth called Aflateen that was developed by Aflatoun Child Savings International in Amsterdam and test-driven by over forty partners around the World.

Intervention: The evaluation experience took place during a ten-month period between December 2012 and October 2013. This paper provides a summary of the context and a justification for the approach.

Research Design: Not applicable.

Data Collection and Analysis: In addition to conventional data collection instruments, we applied participatory inquiry tools from Social Analysis Systems (www.sas2.net) as a means of engaging stakeholders in real-time data collection and analysis.

Findings: Examples are provided to demonstrate how some developmental evaluation findings informed organizational strategic and operational decisions during the life of the evaluation.

Keywords: *utilization-focused evaluation; developmental evaluation organizational learning; participatory inquiry; social and financial education.*

Introduction

In this paper we describe a practical experience where we combined Utilization-Focused Evaluation (UFE) with Developmental Evaluation (DE). This combination is referred to as Utilization-focused Developmental Evaluation (UFDE) (Patton, 2008). The process we describe was part of a consultancy for an international non-governmental organization based in the Netherlands. The work took place between December 2012 and October 2013 with over 50 person days spread over the period. A feature of this article is that it is co-written by the evaluators and the client (the Aflatoun Secretariat) in an attempt to document and reflect from the shared learning.

A note on style: Since this is co-written by three people, the evaluation consultants represented by the first two authors are referred to as ‘the evaluators’; the organization that was the subject of the evaluation represented by the third author is referred to as ‘the client’; when the statements and views are shared by the three co-authors we use ‘we’.

On the surface developmental evaluation (DE) sounds like a common sense approach to track and learn from work-in-progress. For organizations that are innovating and creating models, it is a means to capture experimental work and harness it. As it turns out DE is challenging to implement, especially because it follows moving targets and the evaluators must balance flexibility with focus. It requires evaluation professionals with facilitation skills and a sense of how capacities emerge and evolve within organizations. Developmental evaluators need to get immersed in the context, to appreciate and accompany the innovation process. At times they need to work as enablers, at other times as peers. Most DE manuals emphasize the use of learning methods that help organizations document their work and course-correct. The work involves selecting the tools to apply, sensing when to modify them, choosing when to underline emerging findings, and when to let the clients make sense of the evidence on their own.

In this article we explore how UFE became the overall framework that guided our DE work. We share the elements of the methodology, we present a précis of the data collection and analysis, and we highlight some of the findings. In closing, we review how the findings and the process were used. Our focus is on making the process transparent and to reflect on key decision making

moments that may be of interest to other evaluators.

This paper is organized as follows:

1. The context
2. Research question and justification for the approach
3. Synopsis of the case study methodology and steps followed
4. Some sample findings and their utilization
5. Methodological reflection
6. Conclusions on approach and process

The Context

Aflatoun Child Savings International (hereafter referred to as Aflatoun) is a non-governmental organization (NGO) with its Secretariat based in Amsterdam, the Netherlands. The mission is to inspire children and youth to socially and economically empower themselves, and on that basis to become agents of change for a better, more equitable world. The organization pursues this through a set of social and financial education (SFE) curricula delivered to over 2 million children around the world in over 100 countries through local implementing partners. Aflatoun’s unique structure seeks to create equitable partnerships in which partners feel ownership over their program. As a result, Aflatoun does not provide direct funding to partners for program implementation but rather creates partnerships based on a shared mission and technical support. The Aflatoun Network has the ambitious goal of reaching 10 million children by 2015¹.

The first Aflatoun curriculum was developed for children of primary school age, and as of 2009 plans began to develop a variation for youth under the name of Aflateen as a result of requests from the implementing partners who saw a need for a program of social and financial education for this age group. The MasterCard Foundation then partnered with Aflatoun to provide financial support for the development, roll out, and evaluation of Aflateen starting in 2010. An initial curriculum was created and then piloted in thirteen countries between 2011 and 2012. Following the successful pilot and revision of the curriculum the project is pursuing an expansion to 50 countries in order to reach over 250,000 youth by the end of 2015.

After the pilot had been completed it was then necessary for Aflatoun to produce a midterm evaluation of the Aflateen program. This case

¹ <http://www.aflatoun.org/story>

study provides an account of that evaluation which involved two distinct levels: an organizational evaluation of the Aflatoun Secretariat's role in the program and a separate evaluation of program implementation in several countries.

At the Secretariat level, the focus was an assessment of the Secretariat's role in the development, rollout, and continued technical support of the program. This was designed as a Utilization Focused Evaluation with DE elements and therefore is the primary focus in this paper. For this evaluation component, five themes were selected: (1) the curriculum development process, (2) the ongoing development of the new Aflateen e-learning platform, (3) the technical support provided to implementing partners, (4) the training approach, and (5) evaluation activities.

At the program implementation level, multiple in-country evaluations were conducted that were then synthesized and reviewed by the evaluators. At the implementation level the evaluation was expected to generate lessons for current and future partners and provide an initial assessment of the efficacy of the program. This level was to address the balance between implementation fidelity and local contextualization, define ways to scale up the program in each country, identify short-term outcomes, and explore potential impact pathways.

Research Question and Justification for the Approach

This paper responds to the following research question: *What are the theoretical complementarities and the practical advantages of combining UFE with DE?* For readers familiar with Patton's 4th edition of *Utilization-focused Evaluation* (Sage, 2008), Step 12 includes a meta-evaluation. We propose that this paper is very much part of such a reflection, as a means for the evaluators to become reflective practitioners (Patton, 1994) and contribute to the mutually reinforcing role of both the theory and the practice of evaluation (Schwandt, 2014).

The approach to the evaluation that was proposed by the client referenced Utilization-Focused Evaluation (UFE), Developmental Evaluation (DE) and realist evaluation. In the bid, the evaluators focused on Utilization-Focused Evaluation as the overall approach, with Developmental Evaluation as a likely subcomponent. From earlier experiences the evaluators had come to the realization that UFE works as a broad decision-making framework that gives structure to the evaluation process (Ramírez & Brodhead, 2013). The evaluators and the client

were also aware of the power of DE as a mechanism to learn about and provide course-correction feedback to a new initiative (Gamble, 2006; McConnell Foundation, 2006; Patton, 1994).

From a theoretical perspective it is important to clarify what UFE and DE share and where they differ. Both are presented as independent of any given evaluation method, as decision-making frameworks: "Design, methods, measures, and analysis depend on the priority questions that will support development of and decision making about an innovation based on the nature and stage of the innovation and the situation in which the innovation takes place" (Patton, 2011: 307). Both are participatory in nature, in the sense that the 'users' become co-designers of the evaluation questions; they are involved in the selection and fine-tuning of evaluation tools, and in some cases in the writing up of reports and recommendations. The extent of participation can differ in the two approaches and is distinct from participatory approaches such as community based participatory research (CBPR) and participatory action research (PAR). In UFE, the 'primary intended users' are engaged in the entire process, but those 'users' do not necessarily reflect the perspectives of beneficiaries and service users, as is often encouraged in participatory evaluation and research (Macaulay et al. 1999). In DE, the reporting back can be quick and iterative, allowing for the use of participatory inquiry tools such as Social Analysis Systems, but again this need not necessarily include direct beneficiaries. However, Patton (2011) emphasizes that other methods can be accommodated, depending on the nature of the evaluation questions. In DE the emphasis is on adaptive learning, on real-time feedback, on flexibility and capturing system dynamics (Gamble, 2010). This makes DE compatible with methods that belong to complex adaptive systems, where emergence of unexpected direction is expected as one intervenes in a dynamic context (Westley et al. 2006).

In the evaluation literature that explores 'use', UFE has been a guiding approach that addresses use from the design stage, as opposed to seeking to maximize use when a report is completed. The emphasis on 'users' taking ownership from the start makes UFE the appropriate framework within which DE can be introduced (Saunders, 2012). In both cases, a careful review of stakeholders is important (Bryson et al., 2011). The importance of UFE as the guiding framework within which to introduce DE has been confirmed by its advocates (Patton, pers.com; Gamble, pers.com).

Developmental evaluation, conducted from a utilization-focused perspective, facilitates ongoing innovation by helping those engaged in a pilot project examine the effects of their actions, shape and formulate hypotheses about what will result from their actions, and test their hypotheses about how to foment change in the face of uncertainty in situations characterized by complexity (Patton, 2011, p. 14).

The success of UFE hinges on whether the findings and processes of the evaluation are actually used (Patton, 2008; 2012). As simple as this sounds, this calls for a commitment by 'primary intended users' to drive the process from beginning to end on the basis of concrete 'uses'. The more users are involved, the likelier they will

have ownership over the process, an understanding of the relevance of the findings, and have a stake in implementing the findings (Liket *et al.*, 2014); this applies to UFE and to UFDE (Patton, 2011). In UFE, the evaluators take on the role of facilitators of learning, as opposed to neutral outsider judges. The two most important premises of UFE are that no evaluation should go forward unless there are users who will actually take action on the information that the evaluation will produce, and that they are involved in the process of the evaluation. Patton's original UFE lists twelve steps; while listed in a linear fashion they work in an iterative manner in practice. Figure 1 summarizes the steps and their inter-relationships (Ramírez & Brodhead, 2013).

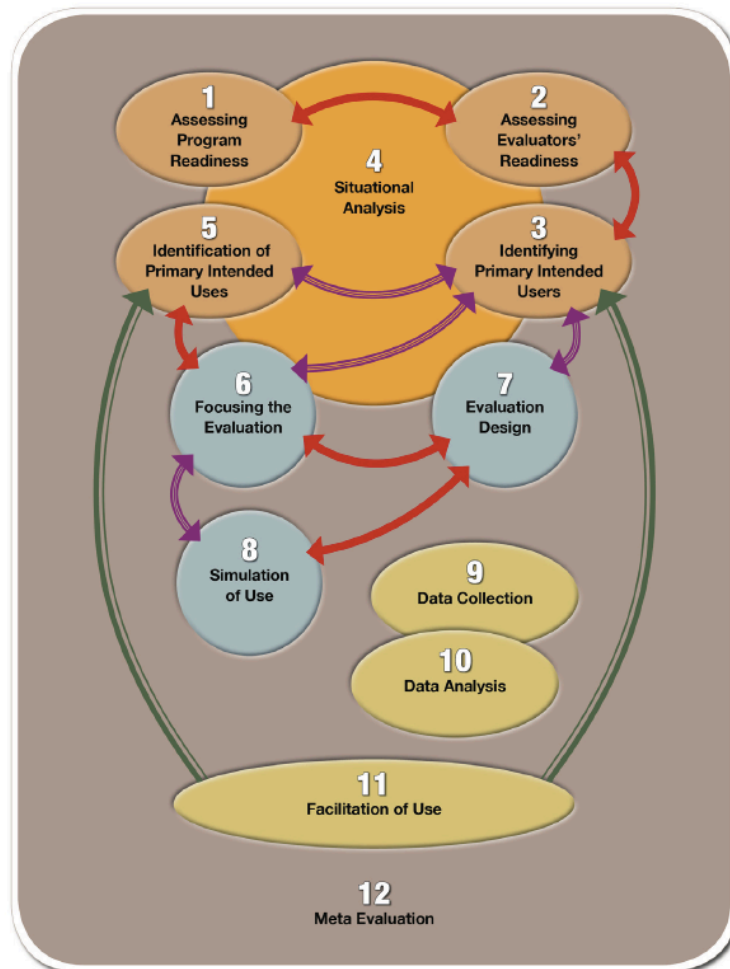


Figure 1. UFE Steps (from Ramírez & Brodhead, 2013)

The first five steps emphasize the need to ascertain readiness of the project, the organization, the funder and the evaluators. They place a focus on the identification of ‘users’ who will drive the process and on the specific ‘uses’ they wish to focus on. Steps 6-8 cover the design phase, where the purposes or ‘uses’ of the evaluation are translated into ‘key evaluation questions’ as a key stage in the process. Step 11 refers to the facilitation of use, where the evaluators assist the client in harnessing the findings and converting them into actions and strategies. This article constitutes Step 12: a meta-evaluation about the experience itself.

UFE has two key advantages: there is continuous attention placed on utilization, hence it is practical. Second, primary ‘users’ take ownership over the design and implementation of the evaluation; in doing this, the process is as important as the findings in shaping future learning mechanisms and creating momentum for implementing the findings. UFE is an approach that is methodologically neutral. In UFE, methods and data collection instruments are selected on the basis of the Key Evaluation Questions (KEQs) that are elicited from the ‘users’. In short, under UFE one can make space for a number of evaluation methodologies including Developmental Evaluation, when the nature of the KEQs calls for that methodology.

In projects, such as Aflateen where field experimentation is underway, a systematic process of learning for course-correction is needed. The literature confirms that DE is particularly relevant for programs involving social innovation (Preskill and Beer, 2012: 5).

Since DE is used as a methodological framework for addressing several key evaluation questions it requires that the evaluators take on a particular orientation including the following roles:

- Seeking a long-term, partnering relationship, with a view to be supportive of an innovative, model-making endeavor;
- Asking ‘evaluative questions’, gathering information systematically, and providing feedback in support of decision making;
- To a certain extent, becoming part of the team; helping to conceptualize, design, and review adaptations; mapping out patterns and relationships; elucidating discussions; exposing assumptions; and facilitating data-based assessments as innovation unfolds;

- Working as insiders-outsiders; co-shaping a process that is unfolding.

A DE process requires *immersion* by the evaluators in the organizational culture, something that will be incomplete but desirable. Within UFE, DE is appropriate as a means “...to develop measures and monitoring mechanisms as goals emerge and evolve. It seeks to capture system dynamics, interdependencies and emerging interconnections.” (Gamble, 2008: 62).

Synopsis of the Case Study Methodology and Steps Followed

The evaluators’ previous experiences with UFE and the broader literature signaled the importance of the primary users taking ownership and control over the process (Liket *et al.*, 2014; Ramírez & Brodhead, 2013). In doing so, they would be assisted in designing a set of practical ‘uses’ and fine-tune the purpose of the evaluation. The ‘users’ were partly self-appointed and partly identified based on the role and position in the organization. The ‘users’ were coached by both the evaluators and the internal research manager in the articulation of key evaluation questions to guide the inquiry. In doing this, the ‘users’ refined what they wanted to learn, how it might be learned, and what would be most salient for use in the next stage of the project. Through this process it became clear that part of the evaluation would be developmental so that a DE approach was integrated.

As noted earlier the evaluation contained two components: a review of performance by the Amsterdam-based Secretariat in the development of the Aflateen curriculum; and a review of the program evaluations about the field testing of the curriculum by partners in different countries. The first component focused on the ‘organizational attributes’ that helped or hindered four dimensions about the curriculum creation and rollout:

- Creation, contextualization and dissemination of the Aflateen curriculum
- Establishment and support of partnerships with organizations to implement Aflateen
- Training of teachers/facilitators/peer-educators to deliver Aflateen
- Evaluation of the Aflateen program

The second component called for the evaluators to

conduct a critical review of program evaluations that were commissioned separately with local evaluators by the client (Aflatoun) and the funder, the synthesis and meta-evaluation of these focused on five questions:

- What is the balance of implementation fidelity and contextualization in the field?
- What are the proximal outcomes regarding each of the five core elements of the curriculum?
- What is the projected causal pathway to long-term impacts and what are the potential barriers?
- What are potential pathways to reaching scale in countries?
- What is the quality of the evidence provided by the partner evaluations?

In combining the UFE with DE the evaluators did not follow all the original UFE steps. The steps followed starting in January 2013 included:

1. Literature review & familiarization
2. First visit to the secretariat
3. Draft interview guide for partners
4. Visit to partners in El Salvador and East and South Africa
5. Second visit to the secretariat
6. Finalization of interview guide and on-line survey
7. Skype interviews and on-line survey implementation
8. Data analysis and preparation of the Interim report
9. Review of program evaluation reports
10. Draft final report
11. Report review and utilization—third secretariat visit
12. Final report

Prior to visiting the Secretariat in Amsterdam, one of the evaluators visited the funder, in this case MasterCard Foundation in Toronto, to confirm their willingness to support the UFE process. Not only did they agree, they stated that, “if the evaluation was useful to the Secretariat, it was useful to them”². Early in the process and after consultation, the funder decided not to become a primary ‘user’, therefore the Secretariat played the

primary active role in determining the direction of the evaluation because all primary ‘users’ were located in the Aflatoun Secretariat. Obtaining a funder’s approval for an evaluation to evolve without their direct control is unusual, and it provided the enabling environment for the type of focused, practical learning evaluation that we detail below. In the evaluation literature, this open attitude by the funder constitutes a major contributor to ‘readiness’ that allows for strategic learning (Williams, 2014).

One of the early steps in UFE includes a situational analysis. This step is meant to help the evaluators appreciate the organizational culture. During the first visit to the Secretariat, the evaluators engaged in a number of tasks that included the understanding of the context and gauging the extent to which the organization was ‘ready’ for this approach. They facilitated the selection of primary ‘users’ among the staff and asked the ‘users’ to consider the specific ‘uses’ for the evaluation. Stakeholder involvement is key in UFE and is important in DE for additional reasons that were relevant in this case: “In a developmental approach, the anticipated and unanticipated must be constantly reconciled. Both beg for adaptive thinking and stakeholder considerations are a fundamental vehicle for adaptation, particularly as the stakeholder themselves may be involved in the program as well as the evaluation” (Bryson *et al.*, 2011, p. 3).

Subsequent to the first visit the evaluators gathered proposed ‘uses’ from the ‘users’, and elaborated a table listing the four topics and an additional topic on e-learning, along with the ‘uses’ and few Key Evaluation Questions (KEQs). From an initial longer list of ‘uses’ and KEQs, they urged the staff ‘users’ to prioritize a smaller set. After several iterations the following ‘uses’ and KEQs were finalized, (see Table 1).

² The MasterCard Foundation hosted a panel during the 2013 Canadian Evaluation Society Conference in Toronto that emphasized learning approaches to evaluation, and one of the authors was invited to present a summary of this evaluation (Ramírez, 2013).

Table 1
Key Evaluation Questions and Uses

| Area / Topic | Use | Key Evaluation Questions |
|----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Curriculum development | Define and document our desired curriculum development process by the end of 2013 for future revisions. | To what extent are partners/ Secretariat satisfied with the process and the final product? |
| 2. E-learning | Create a plan to scale up and broaden the user base of the e-learning. | Did the Implementation match with Strategic Objectives? In other words, did the outputs meet the plan for reaching 50,000 participants? |
| 3. Partners/ Dissemination | To align Aflateen partner support practices with strategic goals and document this for future PM staff. | What are the factors used by PMs to prioritize support provision to Aflateen partners? |
| 4. Training | To differentiate training process for Aflateen vs. Aflatoun as needed. | To what extent do/should we differentiate Aflateen and Aflatoun trainings? |
| 5. Evaluation | To improve a methodology for research calls that enhances learning of partners for the purpose of improving their program and the quality of their monitoring and evaluation. | To what extent did the research call create a learning opportunity for partners? |

Subsequently, for data collection, the evaluators added three columns to the table: “Data Needed to respond to the KEQs”; “Potential Data Sources”, and “Potential Data Collection tools”. This helped narrow down the data collection steps and tools that included: a review of documentation, three visits to the Secretariat (January, March and September, 2013), two field visit reports (where in depth interviews were conducted with six partners), Skype interviews with 9 partners, and an on-line survey (19 English, 11 Spanish; 30/42 = 71.4%).

The evaluators ensured that primary ‘users’ were engaged in the design and approval of all data collection and analysis instruments, though they did establish deadlines for feedback to keep to the schedule and a majority of the feedback was communicated through the client’s research

manager. This included an interview guide and an on-line survey questionnaire. The majority of data collection tools and reporting were in English; though some data collection was done in Spanish and French. Partner visits and in person interviews were also conducted in El Salvador and Uganda, the latter taking place with the participation of five partners representing Botswana, Kenya, Namibia, Uganda and Zimbabwe.

For the Secretariat data collection with staff, the evaluators applied a series of participatory inquiry tools that combine data collection with collaborative analysis (namely from Social Analysis Systems www.sas2.net). A sampling of the tools used is outlined in Table 2.

Table 2
Sample of Participatory Inquiry Tools Used

| SAS Tool | Purpose |
|----------------------------|---------------------------------------------------------------------------------------------------|
| Timeline | To document the stages in the development of the e-learning platform. |
| Force-field | To document the factors contributing to or hindering the effectiveness of the e-platform. |
| Theory of change | To verify the expected trajectory of change of both the overall program and some of its elements. |
| Contribution & feasibility | To visualize the different expectations about reaching outcomes. |
| Domain analysis | To chart the different evaluation strategies along a continuum of purposes. |

The main contribution of the SAS tools was that data was visualized and analyzed with the informants immediately (the informants were staff and/or groups of primary ‘users’). Such immediate feedback aligned well with the emphasis on learning and using findings in an iterative manner in DE.

At the end of the evaluators’ first visit to the Secretariat they made a presentation to the staff summarizing their understanding of the organization, its mission, the scope of the evaluation, and their take on the overall theory of change. This was summarized into an inception report. They provided the client with reports following the other two trips to the Secretariat as well as after the two field visits. Among these reports, the second Secretariat visit report is worth highlighting in that it provided the summaries of the participatory inquiry sessions held with the Secretariat in Amsterdam; feedback and reviews

were received from the ‘users’ and incorporated (some examples are included in Section 4). They produced summaries of the six in-depth face-to-face interviews and of the nine Skype interviews with partners. In order to maintain the anonymity of the respondents, they removed all the identifiable information and provided the Secretariat an anonymized version of the summaries. They used open coding to trawl for patterns in the data. The on-line Survey Monkey™ provided tallies of the 30 respondents, which were shared with the client. They produced a summary template for the program evaluation reports. The client commented and revised these summaries twice, the second round with the assistance of an additional external reviewer. Table 3 provides a summary in chronological order.

Table 3:
Data Collection Tools, Analysis and Client Engagement

| Data collection & analysis | Analysis | Shared with client | Client inputs | Remarks/ UFE Step |
|-------------------------------------------|------------------------------------------|---------------------------|--------------------------------------------------------------------------------------|------------------------------------------------------------------|
| Documentation review | Table summary | No | No | Part of situational analysis (Step 4) |
| January visit to Secretariat report | Report covering early steps of UFE | Yes (a deliverable) | Client follow-up: key evaluation questions (KEQs) | Consultants helped prioritize KEQs (Step 6) |
| February visit to El Salvador report | Focus was testing data collection tools | Yes (a deliverable) | | |
| March visit to Secretariat report | Based on SAS participatory inquiry tools | Yes (a deliverable) | Revisions received and incorporated | SAS tools allow for joint analysis with informants |
| March visit to Uganda report | | Yes (a deliverable) | Comments received | |
| In-depth interviews summaries | | Yes (anonymized) | Interview guides revised with client inputs (several rounds) | |
| Skype™ interviews summaries | | Yes (anonymized) | | |
| On-line survey tally | Summaries produced by Survey Monkey™ | Yes (anonymized) | Revision of some analysis and interpretation | |
| June Interim report | Draft Table of Contents shared | Yes (a deliverable) | Revisions received and incorporated; “Recommendation domains” drafted by consultants | First integration of all findings except for program evaluations |
| July-September Program evaluation reports | Summary format produced first | Yes (a deliverable) | Two rounds of revisions | |
| September Final report draft | Draft Table of Contents shared | Yes (a deliverable) | Drafted and revised prior to last visit; Recommendations drafted by client | Ensured facilitation of use (Step 11) |
| September visit to Secretariat report | | Yes (a deliverable) | Iterative creation of recommendations and implementation plans | Ensured facilitation of use (Step 11) |
| October Final report | | Yes (a deliverable) | Revisions received and incorporated | |

From a UFDE perspective, the evaluators experimented in terms of producing 'recommendation domains' based on their analysis of the different sources of evidence. The intent was to signal trends and provoke a reflection by the 'users'. The user groups then drafted recommendations that were reviewed during the third visit to the Secretariat, with particular attention to practical implementation and utilization. As with other experiences in UFE and DE, some findings were incorporated into new practices while the evaluation was still in progress. In contrast, other findings called for discussion and debate about the implications for course correction and resource allocation. In total, 16 recommendations were drafted jointly by the 'users', and the client's management team. Six months after the evaluation's end, 12 of those recommendations had either been completed or were on-going. Of those implemented, nearly half of them had begun before the end of the evaluation (N = 5; 41.7%) (Shephard & Bailey, 2014, p. 8).

Some Sample Findings and Their Utilization

In UFE users are closely involved in the design of the evaluation. On the other hand, the degree of involvement in data collection, analysis and report writing depends on their availability and the extent to which these tasks are assigned to the evaluator. In DE, the emphasis is on feeding back findings in a structured, and iterative manner, to enable users to adjust their work on the basis of evidence. In our case, the Interim Report became the main communication tool. The evaluators produced summaries of all data collected from the Secretariat and from partners, and they produced 'recommendation domains'. The intent was to show probable recommendation directions on the basis of the 'uses', the key evaluation questions, and the emerging evidence. As has been reported elsewhere, this is part of the art of the 'nudge' in DE where evaluators find opportunities to focus

attention without imposing conclusions (Langlois *et al.*, 2013). Below we share some examples of findings based on the use of participatory inquiry tools during the second visit to the Secretariat. The subsequent examples focus on the UFDE process used in the Secretariat evaluation.

Example 1: One of the five Secretariat evaluation themes focused on curriculum development. The 'use' was to: "Define and document our desired curriculum development process by the end of 2013 for future revisions." The key evaluation question was: "To what extent are partners / Secretariat satisfied with the process and the final product?" To answer the question, the evaluators drew a timeline in the Aflateen curriculum development process; they mapped the stakeholders on the basis of their power to change the curriculum and their actual level of participation, and they listed what needed to be kept, changed, or added to the process.

Each activity led to a diagram, a discussion, and a summary of what this means for the work of the Secretariat. The value of data summaries that are visual and instant is significant. It forces participants to pause and communicate what they have done in a structured manner that allows for instant cross-validation and the identification of points of convergence and divergence. Table 4 is an example about what staff felt could be kept, added, or removed; having this summary on the wall was the basis of a group analysis and discussion. A point of initial divergence was emphasized in which some staff members felt that email consultations were not useful while others felt that they should be kept. This led to a discussion that converged around a consensus that email could be used in select cases, especially to correspond with those who attended a face-to-face consultation, but should be succinct and should not include large sections of curriculum content.

Table 4
Summary chart of What to Keep, Add and Remove/Change in the Curriculum Development Process
(Source: Second Visit Report)

| Keep | Add | Remove/Change |
|---------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| Consult different people (partners, youth, experts). | Regional consultation. | Mechanism to focus content (too broad). |
| Partner workshops. | Consult teachers. | Too long a time line (partners weren't aware of activities). |
| Document feedback. | Consult topical specialists (health, religious leaders, finance specialists, etc.). | Starting a curriculum without fully dedicated workshop with at least 15 partners and experts. |
| Pilot & revision. | Final proofread before layout. | Thorough revision before print & rollout. |
| Limit pilot countries. | Consult youth. | Sending complete curriculum for comments. |
| Using existing materials/resources. | Practice lessons. | Stop using email for consultation (though good feedback from those who had been at workshop). |
| 1 st consultation via email worked well, more so with those who had attended workshop. | Direct pre-testing with youth. | More internal staff feedback. |
| Central set (of materials). | Group simulation games 'make them simple'. | Do not consult at holiday time. |
| Link feedback with face-to-face events. | More curriculum pilot research. | |

When asked what this meant, the staff in the curriculum group reviewed all the evidence from the participatory inquiry tools and arrived at four initial reflections that were later revised during the final visit to create the following three recommendations. Six months later, the first two had started and the last of these had been completed.

- Develop a Secretariat strategy to support country-level curriculum integration, including national events and a guide for partners
- Modify and simplify the curriculum Contextualization Manual through the creation of simple checklists, a list of most frequent adaptations, and FAQs, to increase use
- Document the process for the development of new curriculum based on experience from Aflateen, Aflatoun and Aflatot as part of the Quality Assurance process

Example 2: Another theme focused on the evolution of the e-learning platform. This was one of the most developmental portions of the evaluation as the platform was under construction when this evaluation began and completed around

the time of the interim report. One of the exercises completed was a review of how much staff felt the efforts of the Secretariat would contribute to reaching 50,000 youth versus how feasible it was. The evaluators completed a first "Contribution and Feasibility" chart³ during the March visit to the Secretariat (prior to the official launch of the platform on July 1st), and the Secretariat repeated this exercise a second time in September. The staff involved with e-learning were asked to rank the extent to which the platform was expected to *contribute* to the goal (low=0; high=10) as well the *feasibility* of implementing the platform (low=0; high=10).

According to the March exercise, the 'contribution' rankings ranged from a high of 8 to a low of 4, while the 'feasibility' ones ranged from a high of 8 to a low of 3. The responses showed a large gap in attitudes among the three staff informants with a trajectory that had potential to move towards the top right-side quadrant of high attitudes towards both contribution and feasibility. When the evaluators asked what factors were conducive to that positive trend, the staff identified:

³ This tool is part of www.sas2.net (Social Analysis Systems).

- The flexibility to change course, potentially integrating different media combinations;
 - The planned creation of a more explicit dissemination strategy;
 - Reaching the target of 200,000 Aflateen paper-based participants is highly probable thus creating a large dissemination channel;
 - The platform's content management system (CMS) allows flexibility for changing / adding content if youth find it boring;
 - The funder is in a learning mode and willing to experiment;
- Program managers can be equipped to explain the platform.

In September the same staff members were asked once again to rank the contribution and feasibility of the efforts of the Secretariat. The circles in Figure 2 represent rankings from March and the addition symbols (+) represent those from September. Each arrow represents a different staff member, with the staff represented by the topmost arrow having two initial opinions in March, including the leftmost outlier (4, 3), indicating the variability of the staff member's initial attitude.

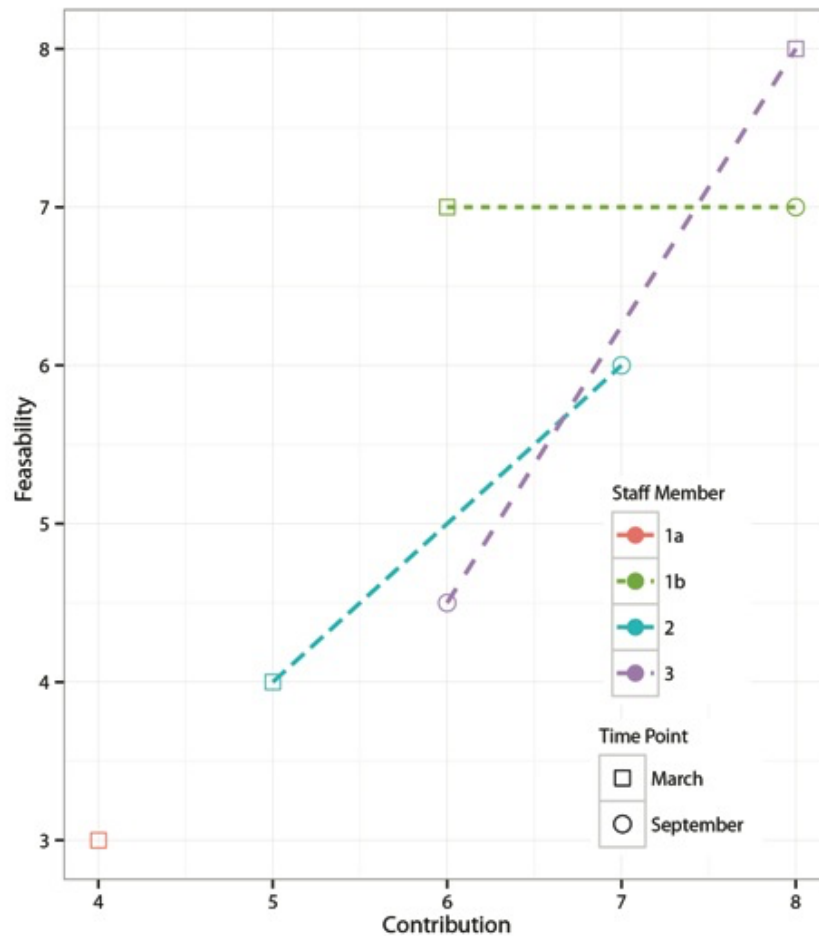


Figure 2. Feasibility and Contribution Evolution

It was evident that the staff members were more pessimistic and critical in March, and had become more optimistic, especially regarding the contribution of the Secretariat's efforts. This is partially due to the successful creation of a platform that has led to the emergence of new partnerships around the world expressing interest in the Aflateen platform because they do not have the internal capacity to build a comparable e-learning system. On the other hand, the most optimistic staff member had a more muted opinion in September. The individual reflected that this was primarily due to the timing of each stage. In March the building stage was almost complete and therefore the process at the time had a bright outlook. However, in September the marketing and dissemination stage was just beginning in earnest highlighting new challenges. Another factor for the less optimistic attitude is

that the skill set for implementation and design of platforms at that moment was stronger than for disseminating a digital product. Looking at the overview of opinions, it was interesting to note that the opinions of the staff had converged and were now much more closely aligned around a cautious optimism.

All of the four recommendations identified for e-learning had begun within the next six months. One of which included a competition during March 2014 for the best social or financial enterprise shared by youth on the site, as voted for by the youth themselves. The winners then become Aflateen Digital Champions who would promote the platform. The competition was intended to expand the user base of the platform and resulted in a large increase in registered users from 368 in February 2014 increasing after the competition to 12,494 (see Figure 3).

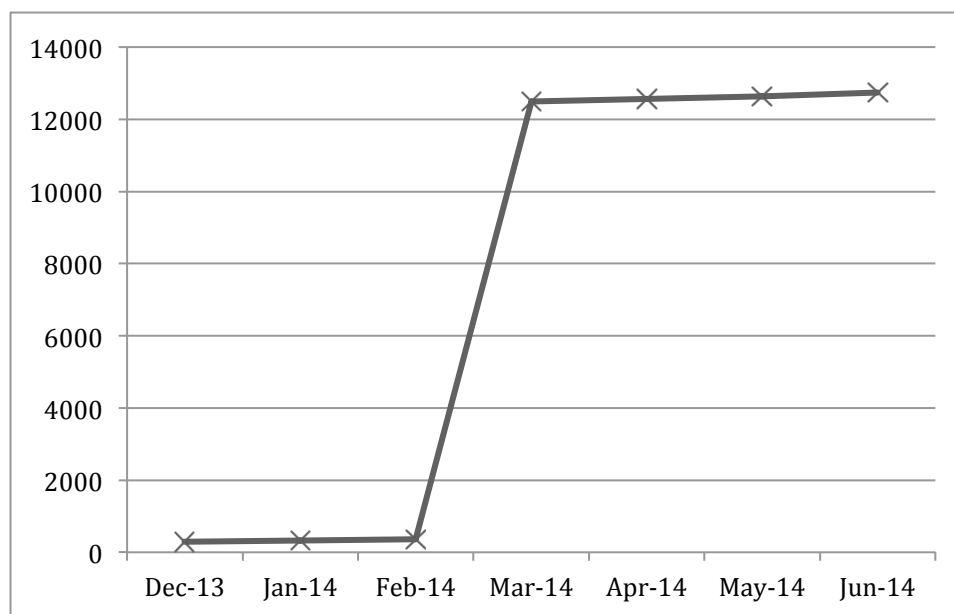


Figure 3. Number of Registered Users on Aflateen Digital

The above two examples are chosen because they fit squarely inside a Developmental Evaluation: they provide insights that the staff can contribute to, own, and use on the spot. They provide evidence of knowledge that is embedded in the collective minds, but often not made explicit and open for group review.

The Interim Report and Nudges

After collecting data from the Secretariat and partners, the evaluators synthesized the findings in The Interim Report and provided "recommendation domains" as a structure for the 'users' to draft recommendations in the final report.

The following example is from the curriculum development theme, the relevant key evaluation questions and use can be found in Table 1.

The findings provided a positive answer to the KEQ regarding the satisfaction of partners and the secretariat with the process and final product. In addition, the proposed improvements address the USE by setting the course for future curriculum revision or development processes.

In addition to the Secretariat conclusions, the following conclusions emerged among partners:

- The curriculum development consultation engaged partners in a meaningful way with close to 77% feeling very positive about the process. Partners also reported benefits that lie over and above the curriculum itself: the participatory teaching methodology was particularly valued; as was the fact that the curriculum can be combined flexibly with other programs.
- A two-tiered curriculum consultation process could work as a mechanism for scaling-up in that higher-level stakeholders in select countries could be involved from the beginning and therefore have a sense of inclusion and ownership.
- The key takeaways in relation to the curriculum development process included the fact that partners were satisfied with the consultation process and that future consultations could be improved by involving local stakeholders, involving more youth and past program recipients, and assessing the feasibility of a two-staged consultation approach. (Adapted from: Interim Report, July 2013)

Using these conclusions the evaluators created the following recommendation domains to nudge the 'users' to reflect on particular trends in the findings as they drafted their own final recommendations for future curriculum consultations processes:

1. The engagement mechanism and the degree of engagement expected from partners.
2. Linking the curriculum consultation processes as mechanisms for scale-up at national and regional levels.

3. The role, purpose, distribution and actual use of manuals.
4. The importance of having financial resources to fund consultations.
5. The framing and communication about the scope of the consultation process.
6. The consultations stages and possible re-allocation of organizational resources within the Secretariat, staff time and un-earmarked budget.
7. Mechanisms to help the members of the Aflateen financial enterprises to improve the odds of selling their products. (Ibid.)

The Recommendations in the final report were written by the primary intended 'users' (during a Secretariat Retreat at the end of July) and revised during the last visit to the Secretariat in September. The September meeting constituted Step 11 in UFE: facilitation of use. By reviewing recommendations that had been drafted by the 'users', and exploring what it would take to implement them, we were jointly able to contribute to refining and prioritizing them.

It also became evident that some findings had already been put to work and that decisions for course-correction had already been taken.

Methodological Reflection

The notion of using UFE as the framework within which to structure DE comes from several sources. Patton (2011) suggests five purposes for DE: i) ongoing development in adapting a project or strategy, ii) adapting effective general principles to a new context, iii) developing a rapid response to a major change, iv) pre-formative development of a potentially scalable innovation, and v) major systems change and cross-scale developmental evaluation (p. 306). In our case we had a combination of i, ii, and iv in that the Aflateen curriculum is a project in the making, it is an adaptation of the Aflatoun curriculum to a new age group, and it seeks to bring the model to a larger scale.

The evaluators sensed a great enthusiasm within the Secretariat for engagement in this process. This was consistent with the participatory nature of DE (Cook, 2006, Preskill & Jones, 2009) and the facilitative role for the evaluators (Langois *et al.*, 2013, Patton, 2011). The openness and support from the funder to allow for the use of UFE and DE was crucial during the evaluation as it allowed the flexibility and freedom to both the

Secretariat and the evaluation to explore new avenues, the importance of which has been highlighted previously (Liket *et al.*, 2014).

An additional dimension worth emphasizing is that the very nature of the changes that were sought are complex, as is the case with many social change projects, especially those with multiple organizations collaborating.

The process and results of collective impact are emergent rather than predetermined, the necessary resources and innovations often already exist but have not yet been recognized, learning is continuous, and adoption happens simultaneously among many different organizations. (Kania & Kramer, 2013, online.)

UFDE responds to this context, especially when integrated with participatory inquiry tools such as those included in Social Analysis Systems (we note that others have flagged this overlap: see Langois *et al.*, 2013). The evaluators also witnessed the value of UFE as a framework that kept the evaluation on track, especially as they have come across other DE experiences where the lack of 'uses and users' as a compass caused uncertainty, especially when new perspectives arose during experimental projects (Leonard, 2013). While we cannot conclude from this study that DE absolutely requires a UFE framework, we have shared evidence of the advantages, namely the importance of users and targeted uses.

Patton concludes that: "...the process of engagement between the primary intended users (social innovators) and the developmental evaluator is as much the method of developmental evaluation as any particular design, methods, and data-collection tools" (2011, p. 335). We the evaluators certainly experienced this as we interacted with the users on a regular basis on the selection of methods, instruments, and analysis. Patton (Ibid.) underlines the importance of rapid feedback, and the SAS tools allowed for that rather well, and the need for a flexible, and adaptive methodology. He adds: "developmental evaluators need to be agile, open, interactive, flexible, observant, and have a high tolerance for ambiguity" (p. 336). The evaluators aimed for this, and perhaps at times both they and the client could have been more flexible, especially when dealing with the review of program evaluations that were contracted separately comprising primarily pre-specified, outcome-focused evaluations, none of which used DE. Lastly, Patton emphasizes that at the heart of DE is a joint

process of synthesis, interpretation and meaning making. The evaluators' innovation with 'recommendation domains' would fit inside this dimension: an attempt to nudge without controlling.

The client advised the evaluators that one of the reasons that they were selected was their experience at capacity development in complex settings, including working in the Canadian north with Aboriginal organizations. Horton *et al.* (2003) argue that capacity development is complex, emergent, and one must exercise flexibility and patience. As it turned out, a great deal of the effort had to do with documenting organizational change, both internally and among partners. The notion of evaluating capacity development was a backdrop to this approach (Lennie & Tacchi, 2013). We feel this process was a joint capacity development effort, where the clients, the funder, and the evaluators were able to create a shared learning space.

Conclusions on Approach and Process

It is appropriate to conclude with the perspective of the client and the perceived challenges and usefulness of the process.

As already noted, the key focus of the funder on learning and usefulness provided the necessary condition to embark on a UFE approach. For the funder, evaluators, and the management team of Aflatoun, it was important to ensure that the approach was clearly understood and mutually embraced. This was facilitated by having the evaluators visit the funder and the Secretariat separately to explain the approach and secure their support.

An important challenge was to ensure that tacit expectations were articulated and incorporated. This included ensuring that the evaluators knew what type of evaluative stance was expected from them for each element of a multi-dimensional evaluation. This challenge surfaced when both the Secretariat and the evaluators had to combine working on the UFDE process during the Secretariat evaluation with the more outcome-oriented approach with the field reports. This tension could have been mitigated if the Secretariat had involved the evaluators more closely in the design and implementation of the field evaluations; however, this was not possible given timelines and resource constraints. The importance of balancing the role of key 'users' as drivers of the evaluation process with the need to include other staff and decision makers also became evident through the process. This was

particularly true once the recommendation and utilization stage was reached and there was a necessary shift from a process driven by a small group of key ‘users’ to broader implementation. This challenge was partly mitigated by the final facilitation of use step of the UFDE process, which involved the entire staff and management team to ensure that recommendations were agreed upon and practical.

Other common evaluation challenges attended the UFDE approach. This included the tension between the status of the evaluators as objective outsiders reporting to a funder and insiders working as facilitators, coaches, and confidants (Mercer, 2007; Parsons, 1974), the challenge of allotting the time required from staff and management to successfully conduct and utilize UFE (Ramírez & Brodhead, 2013), and the importance of having evaluation champions within the organization.

The approach was universally appreciated by staff and management at the Secretariat, as being a useful exercise that helped systematically uncover and address challenges and opportunities. This improved the trajectory of not only the Aflateen program but also other Aflatoun programs. The iterative process and the engagement of staff as ‘users’ facilitated the emergence of novel conceptions of how Aflatoun works with partners and differentiates between multiple curriculum products.

We highlight the following lessons learned. UFE provided the broad decision making framework within which the evaluators and the client confirmed their ‘readiness’ for the evaluation. As expected, the key evaluation questions focused on the piloting of the Aflateen curriculum, which called for a Developmental Evaluation for several of the Secretariat level evaluation questions. The ‘users’ and ‘uses’ set the direction within which the DE was implemented; this allowed the evaluators and the client to have a clear direction. The use of participatory inquiry tools (namely from Social Analysis Systems) allowed for collective analysis of findings on a timely basis. Since the ‘users’ and other staff colleagues were also the informants, the exposure to findings was immediate. This aspect of DE contrasts with the sequence of UFE steps described in Figure 1 in that it adds immediate feedback loops. However, the Evaluation Report was produced and reviewed following the UFE steps. What was unique was the active role by the ‘users’ in drafting the recommendations; this means that Step 11 in UFE (facilitation of use) was very productive as the focus was on prioritizing

recommendations and exploring what was needed to implement them.

Although it is difficult to determine which use cases were initiated by this UFDE experience and which would have occurred without it, it is certain that the process facilitated the application of evidence to all of the five pre-specified themes and even enhanced processes that were not explicitly articulated as ‘uses’ in the evaluation. One of the most lasting impacts of the experience has been the increased clarity with which the Aflatoun staff conceptualizes the discrete curriculum products, such as Aflateen, and how these relate to the Aflatoun Network of partners. The need for more differentiation of support to partners based on their programs was noted as well as the importance of shifting from a network where most interaction flows through the Secretariat to one where partners interact directly with others implementing the same curriculum.

The process was a shared learning experience and demonstrably useful. As evaluators and practitioners we strongly confirm the principle that no evaluation should go forward unless it will be used. We hope our shared experience assists other practitioners and evaluators in conducting practical evaluation work.

Acknowledgments

We are thankful to the staff of Aflateen Secretariat for their commitment and dedication to this evaluation, and to the MasterCard Foundation for their flexibility and openness to a learning approach to evaluation.

References

- Bryson, J.M.; Patton, M.Q., & Bowman, R.A. (2011). Working with evaluation stakeholders: A rationale, step-wise approach and toolkit. *Evaluation and program planning 34*: 1-2
- Cook, T. (2006). Collaborative action research within developmental evaluation. *Evaluation 12(4)*: 418-436
- Gamble, J. (2008). *A developmental evaluation primer*. Montreal: The J.W. McConnell Family Foundation.
- Gamble, J. (2006). *Emerging learning about developmental evaluation*. Hampton, NB, Canada: Imprint.
- Horton, D.; Alexaki, A.; Bennett-Lartey, S.; Brice, K.N.; Campilan, D.; Carden, F.; Silva, J. de Souza; Duong, L.T.; Khadar, I.; Maestrey, Boza, A.; Kayes Muniruzzaman, I.; Perez, J.;

- Somarrriba, Chang, M.; Vernooy, R. and Watts, J. (2003). *Evaluating capacity development: Experiences from research and development organizations around the world*. The Hague, Wageningen, Ottawa, ISNAR, CTA, IDRC.
- Kania, J. & Kramer, M. (2013). Embracing emergence: How collective impact addresses complexity. *Stanford Social Innovation Review*. Jan 21 blog.
http://www.ssireview.org/blog/entry/embracing_emergence_how_collective_impact_addresses_complexity (accessed 28 Aug. '13)
- Langlois, M.; Blanchet-Cohen, N. & Beer, N. (2013). The art of the nudge: Five practices in developmental evaluation. *The Canadian Journal of Program Evaluation* 27(2): 39-59.
- Lennie, J. & Tacchi, J. (2013). *Evaluating communication for development: A framework for social change*. Abingdon, Oxon, UK & New York: Routledge.
- Leonard, D. (2013). Negotiating evaluation design in developmental evaluation: An emerging framework for shared decision-making. Panel presentation. *Canadian Evaluation Society Conference Evaluation Across Boundaries*, Toronto, June 9-12
- Liket, K.; Rey-Garcia, M.; & Maas, K. (2014). Why aren't evaluations working and what to do about it: A framework for negotiating meaningful evaluation in nonprofits. *American Journal of Evaluation* 35(2): 171-188.
- Macaulay A. C.; Commanda, L.E.; Freeman, W.L.; Gibson, N.; McCabe M.L. & Robbins, C.M. (1999). Participatory research maximises community and lay involvement. *BMJ* 1999; 319:774
<http://www.bmj.com/content/319/7212/774>
- McConnell Foundation. (2006). *Sustaining social innovation: Developmental Evaluation*. Montreal: The J.W. McConnell Family Foundation.
- Mercer, J. (2007). The challenges of insider research in educational institutions: wielding a double edged sword and resolving delicate dilemmas. *Oxford Review of Education*, 33(1), 1-17.
- Parsons, H. M.(1974). What Happened at Hawthorne?: New evidence suggests the Hawthorne effect resulted from operant reinforcement contingencies. *Science*, 183(4128), 922-932.
- Patton, M.Q. (2012). *Essentials of utilization-focused evaluation*. Sage.
- Patton, M.Q. 2011. *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York and London: The Guildford Press.
- Patton, M.Q. (2008). *Utilization-focused evaluation*. Sage.
- Patton, M.Q. (1994). Developmental evaluation. *Evaluation practice* 15(3): 311-319
- Preskill, H.S. & Beer, T. (2012). *Evaluating social innovation*. Boston and Washington, DC: FSG and Center for Evaluation Innovation.
- Preskill, H.S. & Jones, N. (2009). *A practical guide to engaging stakeholders in developing evaluation questions*. Robert Wood Johnson Foundation Series. Princeton, NJ.
- Ramírez, R. (2013). Exploring the “L” in evaluation: Putting the learning at the centre. Member of the panel session, the *Canadian Evaluation Society Conference*, Toronto 10-12 June.
- Ramírez, R. & Brodhead, D.P. (2013). *Utilization focused evaluation: A primer for evaluators*. Penang: Southbound.
- Saunders, M. (2012). The use and usability of evaluation outputs: A social practice approach. *Evaluation* 18(4): 421-436.
- Schwandt, T. (2014). On the mutually informing relationship between practice and theory in evaluation. *American Journal of Evaluation* 35(2): 231-236.
- Shephard, D. & Bailey, S. (2014). Combining utilization-focused evaluation and developmental evaluation: The case of the Aflateen Social and Financial Education Program for Youth. Case study presentation at the *CDI Conference: Improving the use of M&E processes and findings*. 20-21 March, Wageningen, the Netherlands.
http://www.managingforimpact.org/sites/default/files/cdi_me_seminar_2014_background_paper_aflatoun.pdf
- Stern, E.; Stame, N.; Mayne, J.; Forss, K.; Davies, R. and Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations*. Report of a study commissioned by the Department of International Development, DFID Working Paper 38. London: DIFD.
- Westley, F.; Zimmerman, B. and Patton, M.Q. (2006). *Getting to maybe: How the world is changed*. Toronto: Random House.
- Williams, A. (2014). *Evaluation for strategic learning: Assessing readiness and results*. Washington, DC: Center for Evaluation Innovation.
<http://www.evaluationinnovation.org/sites/default/files/Williams%20Strategic%20Learning.pdf>