4FRONT

DEVELOPMENTAL EVALUATION OF BUSINESS WITH IMPACT (BEAM) PROGRAMME

WP1: Evaluability analysis

Deliverable 1.1: State-of-the-art analysis

Revised 17.11.2015

Contents

FOREWORD **4**

1 STATE-OF-THE-ART ANALYSIS **5**
1.1 DEVELOPMENTAL EVALUATION AS A TOOL FOR EVALUATING COMPLEXITY 5
1.1.1 Background; the need for a new evaluation paradigm *5*
1.1.2 The Developmental Evaluation approach *7*
1.1.3 How does Developmental Evaluation differ from other evaluation approaches? *8*
1.1.4 When to choose developmental evaluation approach *11*
1.2 TAKEAWAYS FOR BEAM PROGRAMME EVALUATION 12
1.2.1 Suitability of the DE approach to BEAM *12*
1.2.2 Critical consideration for applying DE in BEAM *12*

2 CONCEPTUAL FRAMEWORK FOR BEAM EVALUATION **14**
2.1 UNDERSTANDING THE CONTEXT AND OBJECTIVES OF THE BEAM PROGRAMME 14
2.2 BEAM EVALUATION SET UP AND APPROACH 15
2.3 ASSESSMENT OF PROJECT IMPACT PATHS 16
2.3.1 Step 1: Analysis of anticipated impact areas, aspects and mechanisms *17*
2.3.2 Step 2. Analysis of impact beneficiaries, contributors and stakeholders *17*
2.3.3 Step 3: Construction of an impact path (logframe) with key monitoring indicators *18*
2.4 ANTICIPATED RISKS AND THEIR MITIGATION 19
2.4.1 Programme risks and ways of mitigating them *19*
2.4.2 Risk related to developmental evaluation *20*

3 UPDATED WORK PLAN FOR THE EVALUATION **22**
3.1 EVALUATION APPROACH 22
3.2 DELIVERY OF TASKS AND METHODS TO BE USED 23
3.2.1 WP1: Evaluability Review *23*
3.2.2 WP2: Meta-analysis and biannual reviews. *25*
3.2.3 WP3: Biannual reviews in 2017-2019. ..... *27*
3.3 EVALUATION SCHEDULE 28

ANNEX 1. LIST OF REFERENCES AND EXTENDED LITERATURE **29**

ANNEX 2. SYNTHESIS TABLE FOR THE CONDUCTION OF WP1 ..... **31**

ANNEX 3: DRAFT LOGICAL MODEL FOR BEAM **33**

List of abbreviations

APS	Annual programme statement
AR	Action research
BEAM	Business with Impact Programme
CA	Contribution analysis
CMM	Office of Conflict Management and Mitigation (USAID)
CSBKE	Creating Sustainable Businesses in the Knowledge Economy
DAC	Development Assistance Committee of OECD
DE	Developmental evaluation
DSA	Dynamic systems approach
EC	European Commission
EE	Emergent evaluation
EEP	Energy and Environment Partnership Programme
FP7	Seventh EU Framework Programme for R&D
I4D	Innovation for development
IPP	Innovation Partnership Programme (Vietnam)
KPI	Key performance indicator
MEL	Monitoring, evaluation and learning
MFS	Sustainable Forest Management Programme
MTR	Mid-term review
NGO	Non-governmental organisation
OECD	Organisation for Economic Cooperation and Development
PE	Process evaluation
RBM	Result-based management
RDI	Research, development and innovation
RE	Realistic evaluation
REM	Realistic evaluation model
RTE	Real-time evaluation
SA	Systems approach
SAFIPA	South African Finland Partnership Programme
SAIS	Southern Africa Innovation Support Programme
SE	Systems evaluation
SI	Social Impact
STIFIMO	Programme of Cooperation in Science, Technology and Innovation between Finland and Mozambique
TANZICT	Information Society and ICT Sector Development Project in Tanzania
Tekes	Finnish Funding Agency for Innovation
TF	Team Finland
ToR	Terms of reference
UN	United Nations
USAID	United States Agency for International Development
WEF	World Economic Forum
WP	Work package

Foreword

The developmental evaluation of Business with Impact – BEAM programme begun 25.9.2015 and this report is the first deliverable of the evaluation team.

The primary objective this report is to present the latest approaches and experiences in the design and utilisation of developmental evaluation in Finland and abroad, and draw lessons and guidelines for the planning of BEAM evaluation. This constitutes the first chapter of the report, named ‘state-of-the-art analysis’.

The second chapter reflects the lessons from state-of-the-art analysis to the conceptual framework of BEAM evaluation. It presents a slightly elaborated version of the evaluation approach and design, and particularly how the perspectives of developmental evaluation are taken into consideration. The third chapter then presents a slightly updated and more detailed work plan for the evaluation, paying particular emphasis to the elaboration of the activities in the next step.

It should be noted that this is a working document. The concepts and plans will be updated during the course of the programme and its evaluation. The report merely sets the direction and describes the approach to be taken in the evaluation. In particular, the later stages of the evaluation plan are likely to be adjusted.

One important objective of the developmental evaluation is to document the progress and the choices made during the course of the programme. In this light, we suggest the interim deliverables (i.e. reports like this) will eventually form parts of, or at least ingredients for, the mid-term evaluation report of the BEAM programme, describing the situation and choices, as they are perceived at each current moment.

Evaluation team, Helsinki 17.11.2015

1 State-of-the-art analysis

Evaluation is a critical component of policy-making and programme/project at all levels of Government. In general terms evaluation allows for the informed design and modification of policies and programmes to increase their effectiveness and efficiency. With accurate and reliable information, evaluation provides programme management team and other interested parties with the means to learn from experience, including the experience of others, and to improve service delivery. It serves the dual function of providing a basis for improving the quality of policy and programming, and a means to verify achievements against intended results or unintended consequences (positive or negative).

Evaluation should provide answers to the two-sided question: "Are we doing the right things, and are we doing things the right way?" With answers in the affirmative or with action plans to respond to areas of weakness, evaluation nurtures political and financial support for appropriate policies and help governments to build a sound knowledge base. Thus evaluation can have a strong advocacy role as well as enhancing the sophistication and quality of institutional performance.

Expectations concerning the role of evaluators vary between different evaluation cultures and the paradigms applied. The evaluator is in turn expected to be a neutral judge, a facilitator, a provider of accountability and sometimes even a problem or conflict solver. (e.g. Albaek 2001) In developmental evaluation, the role of the evaluator differs from traditional evaluations. The following sections describe the key characteristics of developmental evaluation approach.

1.1 Developmental evaluation as a tool for evaluating complexity

1.1.1 Background; the need for a new evaluation paradigm

Although, most governments, NGOs, public policy expert, think tanks, industry and development aid organisations agree that evaluation is needed to give critical feedback for policy makers and donors, they also see that evaluation needs a new paradigm and methodology to increase the utilisation of its results. There has been two major external challenges that have accelerated this process: 1) Urgent need for understanding the effectiveness of public interventions or development aid and 2) increasing complexity and interconnectedness in the world and a clear need for policy coherence.

Increasing importance of Results Based Management

In development aid community there has long been a need for increasing aid effectiveness and thus justify the support given to the developing countries. OECD / DAC, UN organisations and many donors has emphasised the importance of Results Based Management (RBM) in development cooperation in the context of *Paris Declaration for Aid Effectiveness (2005)*[1] and *Busan Partnership Agreement (2011)*[2]. Also MFA has published Results Based Management guidelines for development cooperation in 2015[3]. The aim of the guideline document is to outline the results chain approach, which Finland is using in its development cooperation. Similarily, there are increasing pressures to assess the impacts of innovation policies, public business subsidies and R&D funding programmes (see e.g. Technopolis Group & Mioir 2012; OECD 2013; Lerner 2009)

[1] For more information, see: http://www.oecd.org/dac/effectiveness/34428351.pdf
[2] For more information, see: http://www.oecd.org/development/effectiveness/busanpartnership.htm
[3] Results Based Management (RBM) in Finland's Development Cooperation – Concepts and Guiding Principles, MFA 2015

Increased complexity highlights the needs for new approaches

Another driver for new evaluation paradigm was the increasing complexity and interconnectedness of policies. Governments across the globe currently face enormous challenges in trying to cope with the increasing turbulence apparent in many policy-fields and with the evolving complexity and interconnectedness of policies. Attempt to curb the negative impacts of climate change provides a textbook example here of the paralysis of politico-administrative systems in Europe. The problem itself is global in nature and displays both direct and indirect linkages to energy, traffic, entrepreneurship, health and tax policy to name but a few such policy areas.

Sector-based administrative structures do not support comprehensive and appropriate horizontal policy-preparation or the use of the multi-policy assets needed to effectively tackle the problem. The existence of multi-level forms of governance and of various players at the different level of the steering system (in this case: global-European-national-regional and local) makes it even more difficult for decision-makers to see 'the big picture' and to make informed decisions which will really have an impact. All this has altered the dynamics of policy-making and set new restrictions on the credibility of traditional democratic governance models. Traditional evaluation models have been inefficient to cope with increasing complexity and systemic development in global politics and development aid.

Limitations of traditional approaches

Typically the evaluations of public interventions (especially in development policy) are based on rationalistic input-output logic models. However, there has been an increasing criticism that RBM and logic model approach in evaluation are not sufficient tools for evaluating public interventions in complex settings (see e.g. Patton 2011; Pawson 2013).

Rationalistic planning frameworks, which embed the causal logic behind actions (from inputs to out-comes and impacts) have been developed and used extensively, especially in the field of international development aid. The logical framework (or 'logframe') approach has been the mainstream tool for planning aid interventions, both at programme and project levels. Logframes provide a simple and useful planning tool and the basis for evaluating projects or programmes. Logic models are grounded in theory-based evaluation and provide a coherent theory on how activities are intended to generate the results.

Theory-based evaluation approaches highlight that a programme may fail for two reasons: either the programme has failed to put the intended activities into operation (implementation failure) or the activities have failed to bring about the desired effects (theory failure). Where the logic of the project/programme is flawed or activities and outputs do not generate the desired results, the project/programme is likely to fail even if the process of implementation is successful. However, in an era of fast paced change, it is not always possible to predict changes in the economic, social and environmental context that impact on the successful achievement of programme objectives. Too often the inflexibility of the logframe approach can limit staff capacity to adapt to emergent trends by holding them accountable to predicted cause and effect rather than accountable for the ability to learn from the use of rigorous evidence analysis in implementation and to adapt to changing circumstances.

1.1.2 The Developmental Evaluation approach

Rationale: Importance of context

In contrast to the traditional approaches, *developmental evaluation*[4] emphases innovation and strategic learning rather than standard outcomes and logic model -based approached discussed earlier in this paper. In this sense it resembles so called Realistic Evaluation Model. Pawson and Tilley developed the first realistic evaluation approach already in 1997. They argued that in order to be useful for decision makers, evaluations need to identify what works in which circumstances and for whom, rather than merely 'does it work?'. The complete realist question is: "What works, for whom, in what respects, to what extent, in what contexts, and how?". In order to answer that question, realist evaluators aim to identify the underlying generative mechanisms that explain 'how' the outcomes were caused and the influence of context. (Pawson and Tilley 1997)

Context is a very crucial term also for Michael Quinn Patton, widely considered as the founder of the current Developmental Evaluation paradigm. In his 2011 textbook he specifies some ideas behind his thinking:

> "Developmental Evaluation supports innovation development **to guide adaptation to emergent and dynamic realities in complex environments**. Innovations can take **the form of new projects, programmes, products, organisational changes, policy reforms, and system interventions.** A complex system is characterised by a large number of interacting and interdependent elements in which there is no central control. Patterns of change emerge from rapid, real time interactions that generate learning, evolution, and development – if one is paying attention and knows how to observe and capture the important and emergent patterns. Complex environments for social interventions and innovations are those in which what to do to solve problems is uncertain and key stakeholders are in conflict about how to proceed." (Patton 2011)

From defined frameworks to dynamic framing

Emergent and dynamic realities in complex environments and interacting and interdependent elements are fundamentally new thoughts in programme evaluation tradition, which normally follows the causal explanatory path and focuses mainly on deviations of the pre-set targets or benchmarks. Although, Patton is not very explicit in explaining the mechanisms that transform or give form for the new emerging elements he hints that these are resulted from complex set of interaction between actors, ideas and competing preferences or issues.

[4] It should be noted that development evaluation and developmental evaluation are not the same thing. Development evaluation is a set of evaluation practices, approaches, models etc. favoured by international organisations such as the World Bank, United Nations or OECD/DAC together with donor communities. Developmental evaluation (DE) on the other hand is a specific approach to understanding the activities of a programme operating in dynamic, novel environments with complex interactions. These two might overlap but are not the equivalent concepts.

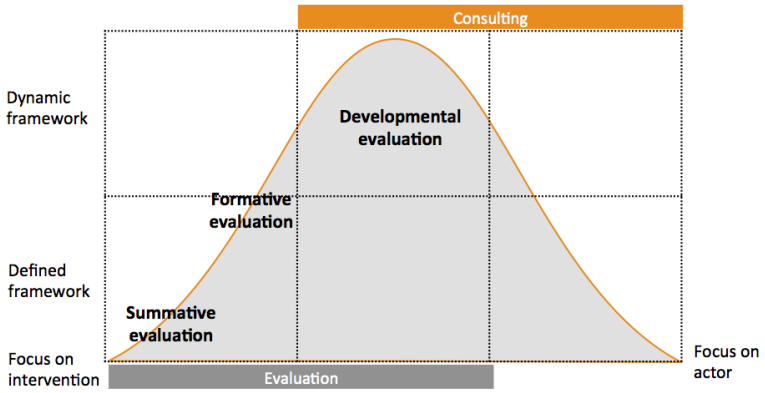

Figure 1. Developmental evaluation as dynamic reframing (Adapted from Patton 2015, 7).

Figure 1 above illustrates how developmental evaluation is especially needed when actions are embedded in a complex system and actors (policy-makers, experts, programme managers etc.) are trying to change it. Actors are making these decisions behind the "veil of ignorance" which most often leads to unintended consequences. Therefore decision-making is constant learning (usually by trial and error methods). Therefore developmental evaluation is process of dynamic reframing. When choosing the "right" evaluation approach the level of complexity and role of the evaluator should be taken into consideration.

1.1.3 How does Developmental Evaluation differ from other evaluation approaches?

Traditional programme evaluation approaches can be categorized roughly into two different approaches: *formative evaluations* and *summative evaluations*. Formative evaluations (or *ex ante evaluations*) are typically conducted before a programme is launched in full-scale. The purpose of formative evaluation is to improve the programme model. Summative evaluations (or *ex post evaluations*) are conducted after the programme (or some phases of it are ended). The purpose of summative evaluation is to assess whether the programme has been successful. Developmental Evaluations differs from both these approaches as it aims to continuously develop the whole process (goals, methods etc) to best respond to the changing conditions.

The main differences between DE and traditional programme evaluation are summarized below:

- **Formative evaluation** aims to *improve and fine-tune* programme
- **Summative evaluation** *tests, proves and validates* programme models
    - Linear problem solving and known cause of the problem, high predictability
    - Programmes are outcome-driven, how best to reach the defined goals
- **DE** aims to *continuously develop the whole process*; both the goals and the methods, to best respond to the changing conditions
    - Nonlinear, complex and dynamic conditions for problem solving, high unpredictability
    - Programmes are driven by the aim to enable systems-change
    - Social innovations and adaptive management
- DE can also lead to generation of a model to be evaluated formatively and summatively.

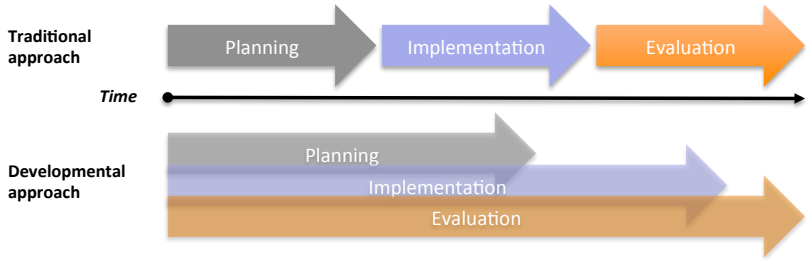

Figure 2. Illustration of differences between traditional (summative) evaluation and developmental evaluation approach. (Adapted from Gamble 2008)

It is important to realise that developmental evaluation is not strictly an antithesis of formative and summative evaluations but mainly an addition to those. As policy-makers most often use a set of policy instruments or so called policy-mix, also evaluator need to adjust to changing need of their clients or substantial changes in a programme or policy that they evaluate. The following table will explain the difference between traditional programme evaluation and developmental evaluation more in details. We have slightly modified the original comparison by Patton (2011) to adjust it to fit into BEAM evaluation framework.

Table 1. Description of the developmental evaluation approach for BEAM (based on Patton 2011)

Evaluation Criteria	Traditional Programme Evaluation	Developmental Evaluation
Purpose and situation	Formative-summative distinction, a priori set linear causal paths, predictable outcomes and impacts, intervention logic well conceptualised.	Supports development in complex, dynamic environment, no known solution to priority problems, multiple pathways possible, iterative learning and exploring new possibilities.
Focus and target	Clearly identified outcomes for intended programme beneficiaries and stakeholders, systems treated as context.	Systems change continuously from small local systems to disruptive cross-scale impacts, provides timely feedback for development and redesign.
Modelling and methods	Design based on cause-effect logic model, counterfactuals a dominant concern, measures performance and success against predetermined goals, evaluators determine the design, which is rigorously methods-focused.	Design the evaluation using systems thinking to capture and map complex systems dynamics and interdependencies, develops measures and tracking mechanisms quickly as outcomes emerge. Evaluator collaborates with those engaged. Co-creating and highly utilisation-focused.
Roles and relationships	Evaluator is independent, accountability externally defined and compliance-based.	Evaluator is a part of the innovation team, a facilitator and learning coach. Accountability centered on innovator´s deep sense of fundamental values and funding priorities and commitment to make a difference.
Evaluation results and impacts	Focuses on validated best practices, detailed formal reports, fear of mistakes and failure, focus on getting (academically) credible evaluation results based on rigorous methods.	Effective principles that can inform practice, rapid, real-time feedback and evidence, building on capacity to challenge, learn and innovate as part of the programme implementation.
Approaches to complexity	Aims for as much certainty and predictability as possible, evaluator(s) attempt to control independently the evaluation process.	Expects uncertainty and unpredictability as givens in complex and dynamic situations, agile learning and continuous interaction with programme management and beneficiaries.

Professional qualities	Methodological competence, analytical and critical thinking, credibility with external authorities, often symbolic values.	Methodological flexibility, adaptability and systems thinking, team work and people skills, able to facilitate rigorous evidence-based reflections to inform action, utilisation-focused approach.

Above we explained how DE relates to summative and formative evaluation approaches. Yet, there is still a big confusion on how other complexity driven, systems oriented or real-time evaluation models relate to DE and what is the major difference between these. Our general observation is that it is difficult to make clear differentiations between DE and other complexity and systems driven models (such as *Systems Evaluation (SE), Real-time Evaluation (RTE), Emergent Evaluation (EE), Realistic Evaluation (RE), Action Research (AR)/ evaluation or Complexity Theory (CT)*). Their boundaries are blurred, approaches overlap and do not neatly fall into different categories by definition. We would like to highlight some of their major characteristics and similarities / dissimilarities with the DE. This comparison has been presented in the Table 2 below.

Table 2. Similarities and dissimilarities between DE and some other recent evaluation approaches.

Evaluation approach	Brief description	Similarities with DE	Differences from DE
Systems approach/ evaluation	Systems approach (SA) is a collection of systemic tools and methods that apply multiple perspectives, study interrelationships and non-linear dynamic changes that alter policies or programmes.	Dynamic systems approach, non-linear and emerging changes. Understanding of the complexity and interdependencies.	Evaluators´ role in supporting policy change not that explicitly stated.
Real time evaluation	Real Time Evaluations (RTE), can affect programming as it happens. This makes it similar to monitoring, and challenges the conventional categorisation of activities as monitoring or evaluation.	Evaluation is integrated with the programme implementation (sometimes also planning). Gives continuous feedback.	It normally follows the causal nature of the programme logic. Does not take programme emergence into account.
Process evaluation	The purpose of process evaluation (PE) is to examine the course and context of a programme. This does not focus on whether or not high-level outcomes have been achieved, but rather on what is happening in the details of the programme itself.	Both emphasise the importance of on-going evaluation and real time feedback.	Process evaluation in normally more goal bound and follows the linear logic of the pre-set targets. Does not pay special attention to complexity, turbulence and interdependencies.
Contribution Analysis	Contribution Analysis (CA) is an approach for assessing causal questions and inferring causality in real-life programme evaluations.	Both are utilisation focused and understand the interrelatedness of simultaneous actions. Both have multi-actor perspective.	CA does explicitly take into account the complexity, emerging elements and recursive effects.
Realistic evaluation	Realistic evaluation (RE) aim to identify the underlying generative mechanisms that explain 'how' the outcomes were caused and the influence of context.	Both have strong emphasis on context and it its influence on changes in programme implementation or policy.	RE is more embedded in the rationalistic causality and linearity as: mechanism + context = outcome

1.1.4 When to choose developmental evaluation approach

It is important to keep in mind that, as Patton (2011) stresses, the Developmental Evaluation approach is not plausible in all cases, rather it should be seen as an approach for specific programmes in specific contexts. Preskill & Beer (2012) have summarised when to choose formative, summative or developmental evaluation approach in the picture presented below.

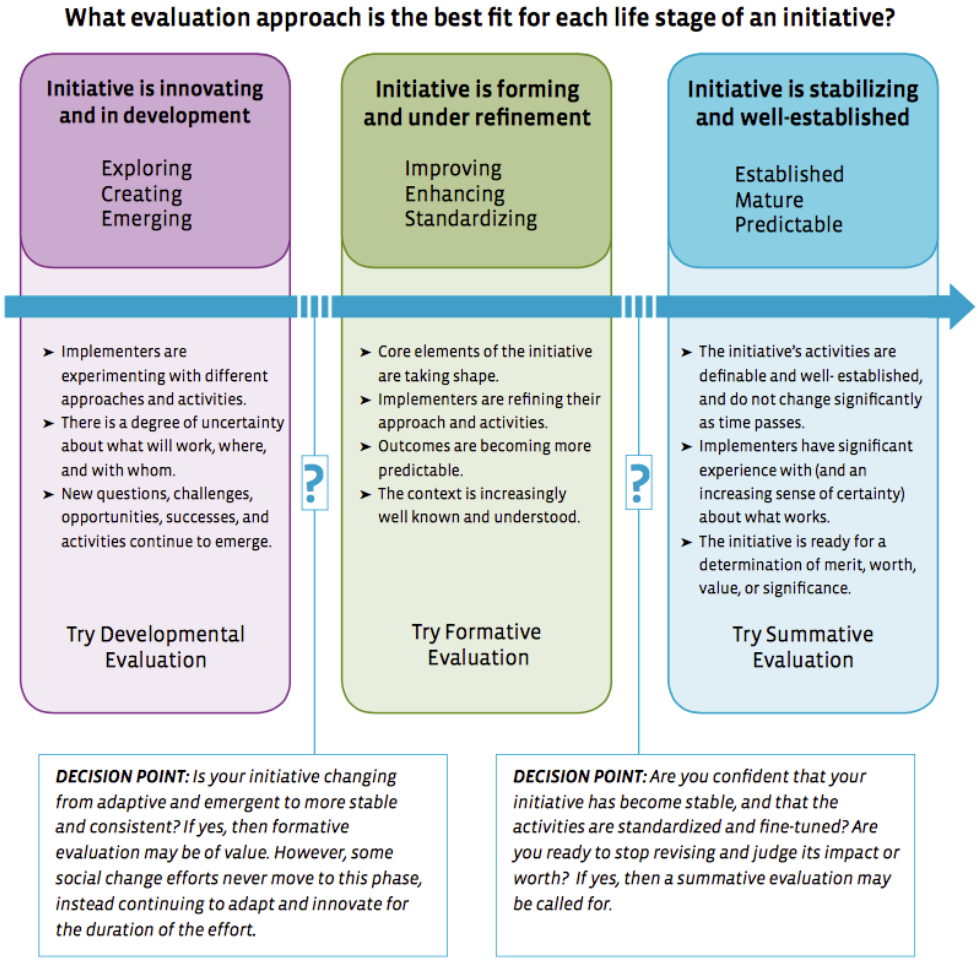

Figure 3. When to choose developmental evaluation approach (Preskil & Beer 2012, 6).

All this can be summarised into the following checklist on when (or whether) to use developmental evaluation:

1. *What is the nature of the problem we're attempting to solve? Is it truly a complex problem?*
2. *What is the system we are trying to affect and how complex is it? (E.g., are there many organisations, actors or activities in this system?)*
3. *To what extent is our intervention a complex or adaptive solution to this problem (e.g., exploratory and flexible, and dependent on the moves or actions of other players)?*
4. *Is our intervention based on a model that is already developed? If yes, do we know what sequence of activities is expected to happen?*
5. *To what extent can we predict most of the short, interim, and long-term outcomes of our intervention? If we cannot predict these, why not?*
6. *Do we need data and feedback as we work to be able to decide next steps?*

1.2 Takeaways for BEAM Programme evaluation

1.2.1 Suitability of the DE approach to BEAM

What then are the key takeaways for applying Developmental Evaluation approach to BEAM programme. Firstly, it should be noted that developmental evaluation is more of an *evaluation philosophy* and a reflective state-of-mind than a compact evaluation approach. Therefore it is impossible to write an evaluation cook-book on DE (and it would be contradictory in terms). In practice, applying DE means re-defining the relationship between consulting and evaluations activities – and setting up practices for continuously adapting and developing the programme, as illustrated below.

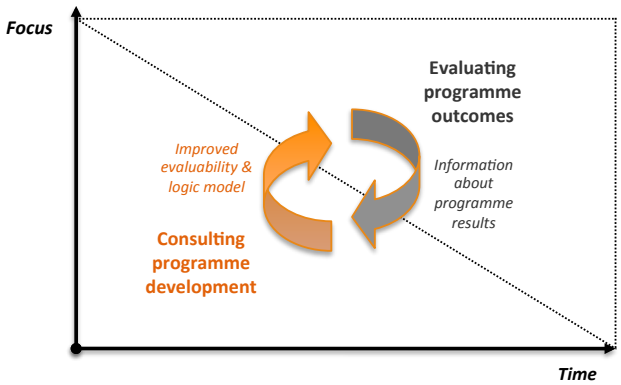

Figure 4. Focus of work in Development Evaluation approach. (Source: 4FRONT)

Secondly, a critical precondition for applying developmental evaluation approach is rather advanced knowledge of evaluation among all parties related to the evaluation process.

Thirdly, DE requires rather mature evaluation culture and capacity to live under uncertainty, which might be a challenge for both decision-makers (in terms of not knowing specifically what kind of impact can be expected) and evaluators (in terms of not having a strictly-defined evaluation framework).

In our view, the BEAM programme appears to meet well all above aspects. Hence, we are convinced that DE is a well-suited approach to the BEAM programme, particularly for the following reasons:

- *It gives long-term directions for the development of BEAM programme*
- *The aim is to achieve transformational changes and DE is suitable for that purpose*
- *It is particularly well-suited to joint or shared programmes, like BEAM*
- *BEAM has multidimensional strategy space (businesses for the developing markets together with poverty reduction goals and human rights).*
- *Both funding organisations (Tekes and MFA) have a strong culture of innovation and a readiness to pilot advanced evaluation approaches*

1.2.2 Critical consideration for applying DE in BEAM

In applying DE approach for the evaluation of BEAM, we highlight the following issues as important or even critical factors that should be well addressed (and to be further elaborated in the Evaluability analysis):

A) Understanding the role and nature of DE in an experimenting programme like BEAM

1. *Understanding what DE is: Which are its strengths and weaknesses, How to use it*
2. *Ability to implement BEAM under a certain level of uncertainty with regard to its ME&L framework*
3. *Acceptance of variable interests, perspectives and dimensions to be integrated in the BEAM*

evaluation.

4. *Willingness for on-going dialogue with DE; openness for questions, advice and also for constructive criticism*
5. *Capacity to utilise DE for rapid assessments and changes to the programme.*

B) Effective utilisation of DE for the purpose of BEAM

6. *DE should allow BEAM to be more explorative, more experimenting and more piloting*
7. *DE should allow BEAM to be able to take more calculated, anticpatory risks*
8. *DE should allow BEAM to be more agile, more quick to change its course and to adapt its processes to new needs*

C) Issues that need to be well addressed and further defined for DE

9. *It is important to define well the mandate of the DE, for example with regard to access to programme data, project specific (confidential) information*
10. *It is important to define well the processes of DE, particularly with respect to BEAM programme management. What are the steps in requesting, delivering and using DE for advice.*
11. *It is important to define well the role and responsibilities of the DE, with respect to BEAM organisations; who does what (Steering Group, Management Team, Coordination Team, Evaluation Steering Group, Evaluation Team).*
12. *The evaluation practices of Tekes and MFA differ from each other. It is important to pick the the necessary and the best practices from both organisations and to synchronise practices where possible. For example, it appears important to adopt from MFA the strong tradition of defining a ME&L framework, the utilisation of logical framework for programme steering, RBM for management and HRBA as a viewpoint.*

D) Collection of data and evidence

13. *It should be clarified / elaborated how the BEAM programme collects data and information from the programme decision-making and implementation, what information is available and how that can be engaged for the use of DE*
14. *It is important to assess how well the information collected, or the information anticipated to be collected, covers the various aspects and stakeholders of BEAM, which are far broader than in normal Tekes programmes*
15. *For the evaluation perspective, it is particularly important to answer the question 'why'; hence to document the reasoning and criteria for selecting certain focus areas, certain partners, certain types of funding, etc. The question of why should then implicitly respond to the question of what kind of impacts are anticipated and prioritised.*

E) Issues that are important for learning and future use of DE

16. *BEAM is first of its kind as a TF programme, but with respect to evaluation, it is also the first programme applying DE. It is therefore important to document and conceptualise the DE process as well as possible, and to pick lessons for further improvements*

2 Conceptual framework for BEAM evaluation

2.1 Understanding the context and objectives of the BEAM Programme

The theoretical approach applied by both MFA (based on OECD /DAC guidelines) and Tekes in their programme impact evaluations follows in general the principles of the Theory of change, applied to a well-defined and targeted public policy intervention (see Fig 5 below). The overall concept and its application are well known and generally approved.

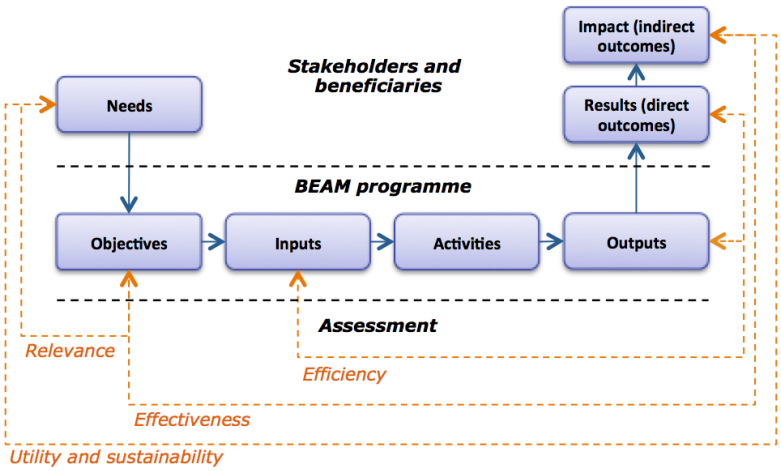

Figure 5. General analytical framework of the evaluation, based on the Theory of change - approach (EC 1997, adapted).

At the same time, the evaluation of the BEAM programme is unique and more challenging than typical Tekes and MFA programmes for several reasons:

- *BEAM is the first programme designed for, and implemented under, the joint Team Finland – umbrella, for which there are common evaluation practices under development. Hence, BEAM is a test-bed for the new TF joint programme evaluation framework.*
- *Although Tekes and MFA have, at large, similar approaches for programme evaluations, they both have different, although mutually complementary objectives, stakeholders and practices for evaluations. This makes the set of objectives and considerations far more broad and complex than typical programme evaluations of either organisation.*
- *By and far, BEAM is the first programme for Tekes and MFA, for which the principles of developmental / on-going evaluation are truly to be applied. Both organisations are familiar with using life-cycle management and result-based management approaches as part of their programme design and implementation, but the full integration of evaluation support to the* ***on-going programme management*** *has not been applied by either to the anticipated extent before. This will not only result changes to the evaluation approach of the programme, but equally will mean developing and testing new practices for the general programme steering and management as well.*

The above set up emphasises not only the broad understanding of the policies and operations of the two organisations behind BEAM, but equally a good understanding the substance topics (RDI, innovation for development) and the nature of collaborating with emerging markets. Solid experience on the various (developmental) evaluation practices and programme management is also of essence.

A particular attention has been paid for the evaluation approach to take into account the **different stakeholder groups of BEAM**; in particular the various stakeholders within the Finnish innovation ecosystems (beyond those directly involved in the project, but likely to benefit from the BEAM opportunities), the direct project partners and stakeholders, as well as the local innovation ecosystem stakeholders in partner / focus countries. Furthermore,

stakeholders within these groups can have different roles, as financiers, co-developers, end users, dissemination partners, etc. This brings an additional aspect to the evaluation, along with the nature of impact in concern (economic, environmental, societal, etc.). To facilitate a sound anticipatory impact assessment at the programme / project inception, we have put special emphasis also on the rationale / relevance aspects, preceding the actual project selection (see illustration in Fig 6 below).

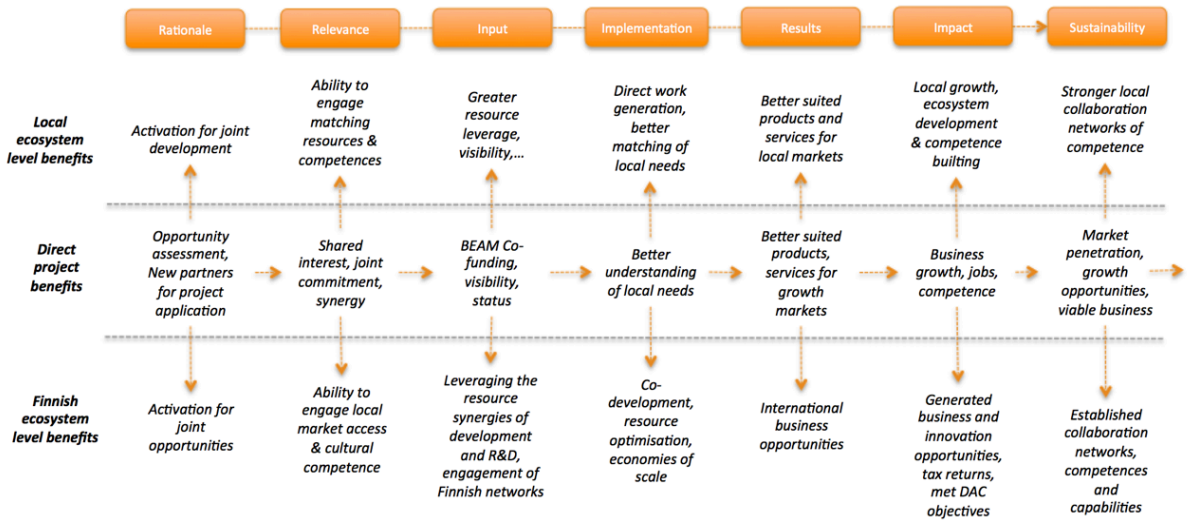

Figure 6. Illustration of BEAM benefits for different stakeholder groups.

With regard to the above, the **typical programme arrangements** of MFA and Tekes differ from each other. In particular, MFA (such as other DAC development programmes in general) has a tradition to use Technical Assistance teams, locally-based partner organisations and often Programme Director's who originate from the 'beneficiary' area – to understand local context and to ensure impact is sustainable. Such approach is not common in Tekes programme, where 'beneficiaries' are normally Finnish stakeholders and coordination / implementation is done mostly home-based. BEAM will be a combination of these approaches and the evaluation (WP1) will aim to distinguish the pros and cons of the programme approach and set up for future learning.

2.2 BEAM evaluation set up and approach

The developmental evaluation of BEAM will be implemented in parallel with the programme during the years of 2015-2019. Developmental evaluation is a particular approach which supports innovation development, guides the adaptation to emergent and dynamic realities in complex environments by continuous utilisation of evaluation results in programme management. This implies analysing and reflecting at any given time the current state of the programme and the decisions at hand, and reflecting these against the overall objectives of the BEAM programme. Typical to programmes that are operating in developing countries is, that the operational conditions and contexts can change rapidly and significantly, which requires a certain amount of agility and adaptability from the programme.

It is important for the evaluation to take into account the **life-cycle stage of the programme**. Hence, in the beginning part of the programme life-cycle, the evaluation will focus more on the baseline and impact logic definition (=monitoring framework with indicators), during the implementation on the impact anticipation and simulation of the different programme choices, as well as on the elaboration of the monitoring and evaluation practices for the programme management. Towards the end of the programme, the evaluation will focus more on the result analyses and impact assessments, as well as on the

sustainabilty measures. During the lifecycle of the programme, also the approaches and subsequent methods, as well as the effort put to the evaluation is deemed to vary (see illustration in Fig 7 below).

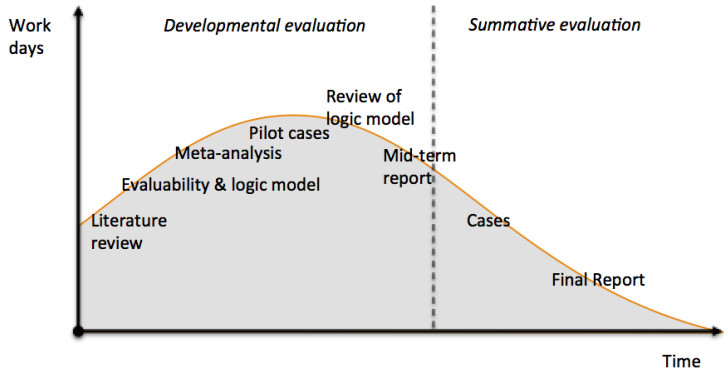

Figure 7. Use of developmental and summative evaluation approaches in BEAM

The evaluation is designed to **continuously support the programme management** by analysing its options and identifying important development trends in the programme's operating context, and when needed, simulating different impact scenarios as proactive choices for the programme decision-making (see Fig 8 below). This will allow the programme to be constantly directed towards the best, or the most optimal anticipated impact paths.

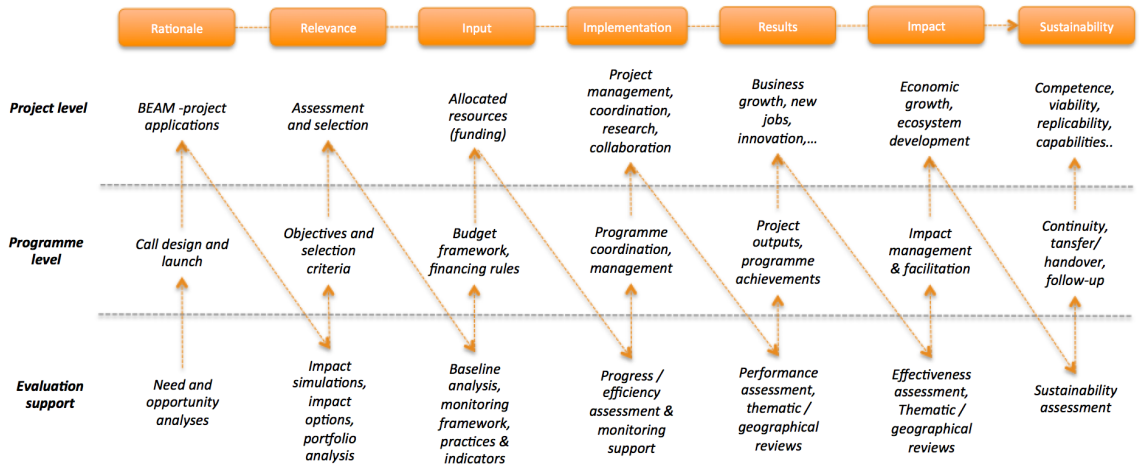

Figure 8. Contribution logic of the developmental evaluation for BEAM programme management and projects

2.3 Assessment of project impact paths

Typical challenge for anticipating the impact of an innovation project is the fact that **innovation is by nature unpredictable**, and it is often impossible in the beginning to set clear definite goals for which to evaluate against. This is particularly the case in innovation for development (I4D) projects, where operational conditions are less predictable than in 'normal' RDI projects. In the big picture, we are essentially looking at multi-faceted systemic changes, and the question is how to measure, document and manage that process. For that, the programme component / result area level is often the most appropriate level (not too broad, not too detailed) to analyse impact paths and to set functional monitoring indicators. Below is our suggestion on how it is to be done in BEAM.

2.3.1 Step 1: Analysis of anticipated impact areas, aspects and mechanisms

The starting point of each case analysis is the various impacts the applicant organisation (in this case the company) expects to generate with the project, as stated in their BEAM application. In the hypothetical example case of testing diabetes measurement system in an innovation platform in India, they could be economic impacts (new jobs through the Indian distribution channel of the system, new jobs in community health care), impacts on governance, participation and security (e.g. women's access to diabetes testing in an environment where due to gender bias women don't have equal access to health care), impacts on access to services and welfare (e.g. better access to diabetes measurement for different social classes, earlier detection, better access to information on diabetes in general, better access to diabetes measurements for rural population), or impacts on capacity development (e.g. new knowledge for medical personnel from doctors to community health care workers, for selected communities new skills on participatory methods, ideation and innovation). The below table presents the attribution of how such impacts are typically generated during the life-cycle of the project.

Table 3. Examples of the different impact aspects of the Diabetes Measurement System and their generation during the life-cycle of the project

Contribution	Rationale	Relevance	Input	Implementation	Results	Impact	Sustainability
Economic impact	Unmet market potential	Market renewal and development	External financial engagement & leverage	Local business & ecosystem development	Better diabetes products and services to local markets, improved sales	New local jobs, export / royalties for Finnish companies	Local ecosystem strengthening, business scaling capabilities, service value chain development
Environmental impact	--	Local services		Less logistics (climate)	Less import, less waste, less energy	Environmental awareness of local solutions	
Societal impact	Unmet diabetes needs	Better health solutions	Inclusive investment opportunity	Local engagement and competence development, local needs recognition	Better diabetes measurement systems & treatments	Health improvements, more equal health, innovation knowhow	Improved collaboration capabilities, Business and innovation competences

For each of the anticipated impact areas, there should be an identifiable impact mechanism, i.e. a realistic and distinguishable indication of how the project will contribute through an unbreakable chain of actions and direct consequences to cause this impact (i.e. impact path). Once this has been identified / constructed, monitoring indicators can defined on the key parts of the path to indicate project progress, performance, result generation etc.

Step 1 outcome: Project impact path, with key impact aspects and related progress / performance indicators.

2.3.2 Step 2. Analysis of impact beneficiaries, contributors and stakeholders

The second part of the project impact analysis focuses on the wider / more indirect impacts of BEAM projects, with emphasis e.g. on beneficiaries and stakeholders in local ecosystems, what the project contributes to them and vice versa. For the generation of wider and sustainable ecosystem effects, project interaction with community/ecosystem change agents is crucial, and for that they need to be identified and engaged at the inception. Qualitative indicators, to describe for example the richness of interactions, are needed, and both the quantity and quality of active contributors in the ecosystem need to be analysed and documented throughout the project. The below figure illustrates how the anticipated benefits are generated during the lifecycle of the Diabetes Measurement System -project to different stakeholder groups.

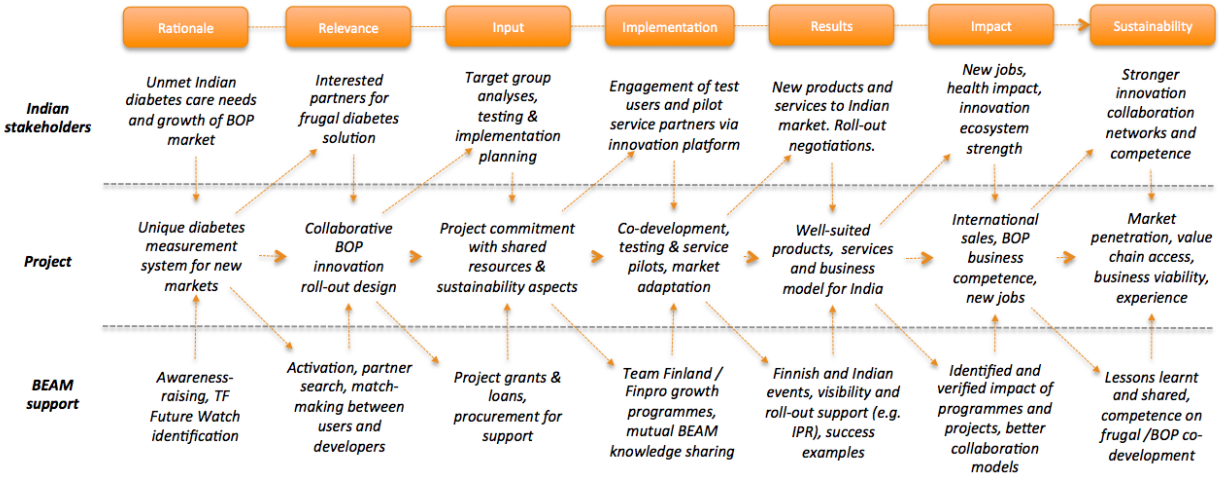

Figure 9. Impact generation path of the Diabetes Measurement System project from stakeholder perspectives.

Participatory methods can be used with a mixed group of stakeholders to investigate the expected impact, explore potential unexpected neutral or negative outcomes, and to create jointly few alternative scenarios for the future use and ecosystem of the product. The scenarios can be used to create roadmaps and to throughout the project to check which path the project is on, what has changed in the assumptions or in the environment, and what changes need to be made in the implementation to reach one of the preferred scenarios and to achieve the impact. The scenarios need to look at sustainability from a business point of view, as well as local cultural aspects influencing business models and usage, such as preference to micro-selling and micro payments, importance of household frugality, importance of relationships and trust, etc.

Scenarios are likely to have different impact profiles. For example, the scenario targeting on rural population and relying on community health workers as a distribution channel via a micro-franchise model could have a stronger job creation impact for women than a scenario where the focus is in fast growing cities and using doctor-driven clinics as the primary channel. At the same time, economic empowerment of women may have unexpected negative outcomes for them, if their spouses or families feel threatened or excluded, or if the surrounding community feels they are stepping beyond the traditionally accepted roles of women. Ideally these risks should be identified at the scenario phase as well, and a risk mitigation plan created. It is important to map out all meaningful and relevant potential impact factors, even though not all of them will be part of the final logframe.

Special attention will be paid to the learning of and impact on the innovation platform and it's stakeholders, inclusion of women and other marginalised groups (i.e. MFA cross-cutting objectives). The role of the evaluators is to analyse the key players and change agents in the ecosystem and to keep a discussion going with them to identify needs for interventions. As Indians culturally prefer not to say ”no” or to give answers which can be seen as disappointing, traditional surveys or questionnaires may not work, and communication strategies need to be tailored for each situation.

Step 2 outcome: Wider impact analysis and with (qualitative) impact indicators.

2.3.3 Step 3: Construction of an impact path (logframe) with key monitoring indicators

Based on the above analyses and resulting scenarios and roadmaps, there will be a concise suggestion for the 3-5 impact KPIs, and how and when they will be documented and

followed up. This allows the management team to get timely information on expected impact, and to see which projects need extra intervention and support to ensure desired results. The KPIs need to be agreed on jointly by the company, as well as by the BEAM management team. A simple visual indicator, for example traffic light red/yellow/green should be shown for each KPI to allow the steering committee and other stakeholders to get the big picture easily. There should be a mechanism for refining the KPIs if and when the goal scenario or the selected road map is modified as typically happens in new ventures and innovation projects. Selected KPIs should ideally be a mix of qualitative and quantitative indicators.

Step 3 outcome: Project impact path with monitoring indicators (KPIs), risk factors, etc = Logical Framework

2.4 Anticipated risks and their mitigation

For the purpose of the developmental evaluation of BEAM, we can categorise anticipated risks into a) risks related to the successful design and conduction of the BEAM programme itself, and b) risks related to the successful design and conduction of the programme evaluation. The two are obviously closely interconnected, but approach the issue from different perspectives. At this stage, the risk assessment is done purely on the basis of evaluation literature and experience of similar programme activities. Hence, this matrix provides a framework for a programme risk analysis, which will be elaborated during the course of the programme and its evaluation.

2.4.1 Programme risks and ways of mitigating them

Table 4. Typical programme risks and their assessment in a risk-management matrix

Type of risk	Possible reasons	Potential consequences	Mitigation measures (DE perspective)
Rationale: Unclear, missing or ill-founded programme strategy. Unclear needs for programme intervention.	Lack of pre-programme analysis, bad reasoning, bad planning, lack of strategic processing & vision. Political compromises.	Inconsistent / random programme results. Unlikely programme success.	Properly studied programme context, sufficient analysis of stakeholder and beneficiary needs and rights, well-thought programme strategy, elaborated programme plan.
Resourcing: Insufficient volume or type of resourcing for the programme requirements.	Shortcomings in programme planning and preparation. Inability to assess the amount & type of resources needed. Too high expectations / ambitions for available resources.	Programme shortcomings in some or all areas. Inability to meet set expectations.	Realistic programming. Scoping and focusing the programme activities only in areas, which can be sufficiently resourced. Reserving resources for programme contingency. On-going assessment of programme efficiency and anticipated impacts of choices.
Organisation: Inappropriate or weak programme organisation, leadership.	Shortcomings in programme planning and preparation. Lack of needed coordination / TA – resources + experts when needed. Unavailability of experience. Lack of good programme governance (Steering and monitoring). Lack of resources devoted to programme organisation.	Inability to understand, assess and make substance-related decisions. Inability to operate efficiently. Inability to execute programme activities in complex situations. In ability to foresee risks and adjustment needs. Inability to make corrective actions and to change programme direction.	On-going assessment of the functioning of the programme team and the Steering Group. Advance planning of specific programme tasks and assignments for possible outsourcing of competence. Preparation for substitutions and ad hoc needs.

Reach and engagement: Inability to reach and engage anticipated stakeholders and beneficiaries.	Shortcomings in programme planning, preparation and organisation. Inappropriate identification and analysis of target groups and stakeholders, and their needs and motivations. Lack of engagement mechanisms. Unclear or insufficient foreseen added value for potential programme stakeholders and beneficiaries.	Inability to carry through activities with anticipated volume and quality, as partners cannot be engaged. Typically shows in lack of participants to events, lack of applications for calls, lack of commitment from programme partners.	Ensuring continuous assessment and feedback from target groups, beneficiaries and stakeholders. Ensuring evidence-based decision-making, as much as possible. Rational and elaborated process for project selections. Continuous portfolio management with foreseen impact assessment.
Implementation: Difficulties (i.e. delays, diversions) in carrying through planned activities.	Shortcomings in programme planning, preparation and organisation. Any or all above reasons behind + inability of the programme organisation to operationalize and deliver activities. Often many practical reasons caused by a new culture and operating environment (language, practices, unforeseen bureaucracy, politics,..). Sudden changes in the operating context, for which the programme is not prepared.	Delays, diversions, rising costs, shortcomings in activities. Slow progress, serious delays or shortcomings. Low participation or low success of events. Diversions from original plans in order to cover for bad planning or to adapt to new situations.	On-site assessment of programme implementation. Direct and immediate feedback mechanism from partners, stakeholders and beneficiaries. Well-thought progress and performance indicators. Process for quick adjustments as needed.
Outputs and outcomes: Inability to deliver results in sufficient volume and quality.	Shortcomings in programme planning, preparation, implementation and particularly in monitoring. Inappropriate /ill-functioning monitoring and evaluation. Programme is focusing too much on activities and on the development work, instead of on the delivery of outputs and generation of outcomes. The activities conducted are not of sufficient quality and practical relevance in order to generate real added value to beneficiaries.	Disconnection between planned activities and desired outcomes. Programme impact remains narrow and benefits focus only to those closely involved and directly participating. No wider impacts, no big changes in the system or new operating practices adopted by the wider community. Programme value remains limited.	Clear definition of anticipated impacts, their target groups and mechanisms delivering the impact.
Sustainability: Inability for the programme results to remain and lessons to be passed over.	Shortcomings in programme planning, preparation and implementation. Learning component missing in monitoring and evaluation. Lack of 'exit plan'. No, or too late planning for programme hand over or continuation mechanisms & process. Programme planning and steering dominated by the donor / one interest partner. Insufficient buy-in from stakeholders and local partners.	Lack of increased capabilities with local partners. Lack of structural changes with beneficiaries. Lack or short of sustainability of the generated impact. No or only partial continuation / hand over. Programme value remains short (in time).	Ensuring programme is designed for delivering structural changes / sustainable benefits. Ensuring stakeholder buy-in and engagement from early on. Agreed plan for gradual handover. Agreed exit / continuation plan.

2.4.2 Risk related to developmental evaluation

Table 5. Potential risks of DE and suggested mitigation measures

Type of evaluation risk	Possible reasons	Mitigation measures
Too loose evaluation framework, bringing about only vague answers how well the	Innovative and explorative programmes can have an opportunistic approach with somewhat	Working out the programme and evaluation framework at the inception of the programme

programme is meeting its targets.	unclear or (ad hoc) changing needs for programme intervention. Unclear, missing or ill-founded programme strategy. Loosely defined programme objectives, activities and targets against which the evaluation should be carried.	evaluation. Testing the framework with pilot cases and further defining the framework on the basis of evaluation feedback and programme experience.
Lack of sufficient background information; documentation, source data and evidence for proper assessment.	Insufficient volume or type of resourcing for the programme requirements.	Systematic collection of programme information from the very beginning (i.e programme justification before its inception). Conducting additional data searches (e.g. baseline information) as needed. Ensuring systematic data gathering and storage during the course of the programme. Monitoring also the data collection, as it is an important factor for learning and improvement.
Unclear role / intervention of evaluators to the programme decision-making and implementation, resulting in loss of integrity or a 'myopia' in evaluators' observations.	The role of evaluators not well defined and elaborated with respect to programme management, coordination and monitoring. The working methods and collaboration practices between the programme management and evaluators not well defined. Evaluators are assigned to assess micro-level issues, losing perspective of the whole programme and its impact.	Clear definition of the role of evaluation with respect to a) access to data and information, b) programme design, c) programme management and monitoring, d) delivery of views and advice for programme steering and management. Detailed distinction between programme management and monitoring tasks (Programme management & coordination team) and programme advice and evaluation tasks (Evaluators).
Unrealistic expectations for the evaluation, in terms of its ability to always ensure 'best programme decisions' or to cover all needed aspects.	Too broad scope / coverage of the evaluation with respect to available resources. Too much 'outsourcing' of programme decision-making to evaluators. Too much reliance on anticipatory impact assessments as a piece of evidence.	Well-designed evaluation plan, with key focus areas and aspects. Not too broad scope. Ensuring the role of evaluators remains as advisors to the programme management.
Slow or ill-functioning process for delivering and using evaluation advice.	Inappropriately elaborated processes for the developmental evaluation to deliver advice to the programme steering and management. Wrong kind or not-timely information and advice delivered to the programme management. Programme steering and management not prepared or committed to utilise DE advice in a 'dynamic way' in programme decision-making.	Cleary defining and designing what kind of information / advice is needed from the Evaluation, at which point, etc. Ensuring the evaluation information / advice is delivered in good time to allow it to be processed and taken into account in planning. Vice versa: ensuring the requests for advice come early enough to allow them being properly processed.
Lack of learning and engagement mechanisms.	Lack of systematic documentation of the programme decisions, activities, results etc. Lack of continuous dialogue of programme direction, performance, added value, etc. Lack of documentation of the evaluation process and the results incurred during the course of the programme Lack of mechanisms to engage the various stakeholders and beneficiaries to learn from the programme during its course.	Defining the learning aspects of the programme and the mechanisms that allow different stakeholders to be part of the learning process. Building a feedback mechanism also for the evaluation advice for learning purpose.

3 Updated work plan for the evaluation

3.1 Evaluation approach

Developmental evaluation approach is typically strongly participatory, supporting the programme management in smaller and bigger decisions throughout the life-cycle of the programme and ensuring rational decisions can be made, based on well-considered information or evidence to bring the programme to its anticipated impact. The developmental evaluation of BEAM has a number of pre-designed tasks (such as mid-term review) and aspects to be carried out. The Figure 10 below illustrates how these elements form the overall framework (Evaluation design) for the evaluation of BEAM.

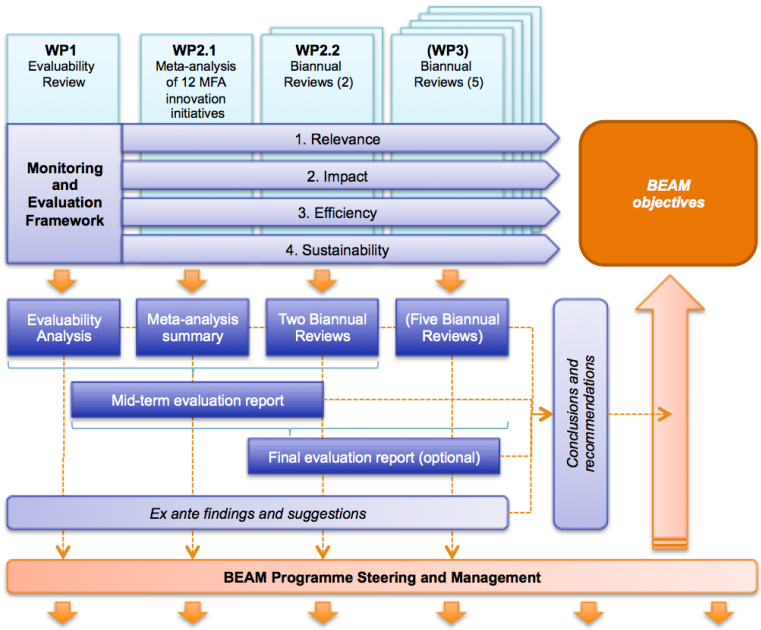

Figure 10. Evaluation design showing specific tasks and their contribution to BEAM objectives

The BEAM programme aims to support Finnish companies, NGOs, research organisations, universities, universities of applied sciences and others in developing, piloting and demonstrating innovations that improve well-being in poor countries, while giving rise to international business opportunities for Finnish companies. The implementation of the programme takes place in a rather complex and highly interdependent environment. This sets certain rather new and challenging characteristics and criteria also to the evaluation design and methodology. Informed by systems thinking and sensitive to complex non-linear dynamics developmental evaluation supports not only the innovation process per se but also reflective and adaptive programme (re)design and management. At best developmental evaluation supports innovation development to guide adaptation to emergent and dynamic realities in complex environment and thus provide real-time evidence for governing the programme and validating its effectiveness track.

3.2 Delivery of tasks and methods to be used

3.2.1 WP1: Evaluability Review

The proper design of the programme and all related considerations of its anticipated impact will be instrumental to the further decisions during the course of the programme and will reflect the overall programme outcomes as well. It is therefore important that sufficient emphasis is put into the good inception of the programme, in the design of the evaluation (MEL) framework and baseline definition, as well as in the setting up the related practices. The whole evaluation team will provide support to this process in several ways.

As its first task, the evaluation is to conduct a short **state-of-the-art analysis** (1.1) of ex-ante and real time evaluation with recommendations for the Tekes programmes. This analysis will include a desktop study of the key approaches in utilising anticipatory and developmental evaluation approaches for programmes by different countries, organisations and funding agencies. The result will present the current status, the key trends, general categorisations / conceptualisation of the different approaches and a discussions of their benefits and usefulness for the BEAM and Tekes purposes. The outcomes are presented and discussed at the Evaluation Steering Committee in November 2015.

The second task of the team will be to carry out an **analysis of the ramp-up phase** (1.2) of 2015-2016 with an emphasis on the planned activities for setting up / delivering the BEAM programme and recommendations for their fine-tuning. This will be conducted by studying in detail the programme documents, background documents and by reflecting these against a) the anticipated evaluation plans and b) the experience of the evaluation team on the successful conduction of similar programmes. This will be complemented with further interviews and discussions, and concluded in an internal workshop for BEAM management & steering. The key inputs for the analysis of the ramp-up phase are

- *Programme descriptions (planning documents) and presentations*
- *Documentation regarding the design of the calls, project selection criteria and the outcomes of selection (project portfolio / project map)*
- *Interviews with key programme partners and stakeholders (BEAM management team, Finpro, MEE, Finnvera, etc)*

The workshop will present simulations of key impact scenarios / discuss the anticipated consequences of strategic programme choices. For practical reasons, it has been suggested that the workshop will be held in connection / jointly with the BEAM management team, in which some three hours are reserved for the discussion on the evaluation aspects. BEAM coordination has proposed tentative dates for this being 1st and 3rd December 2015.

In that **workshop**, the following three issues should be presented and discussed:

1. *Key messages from the state-of-the-art analysis (i.e. what is specific to the developmental evaluation of BEAM) and findings from first interviews*

2. *A 'reconstruction' of BEAM programme for the purpose of the evaluation – the key elements as they are seen from the evaluators' perspective (in a logical framework)*

3. *A draft framework for result-based management (RBM) of BEAM; for discussion.*

The formal deliverable of the analysis of the ramp-up phase (1.2) will be the presentations at the December workshop. The substance outcomes and feedback from the workshop will be integrated into the report of the evaluability analysis (1.3) due in January 2016.

The most important task of the first work package is related to analysing, and through that,

ensuring the **evaluability of the BEAM programme** (1.3). It will ensure the programme objectives are clearly defined, will follow a concrete change logic and serve the programme purposes. As a result, it will establish a result-based monitoring, evaluation and learning framework (MEL) for BEAM.

The figure 11 below indicates typical decision points during the course of a programme. The highlighted decision points illustrate a hypothetical selection, for which the developmental evaluation should connect with or pay particular attention to. Such anticipated **key decision points are to be identified** for BEAM during the evaluability analysis.

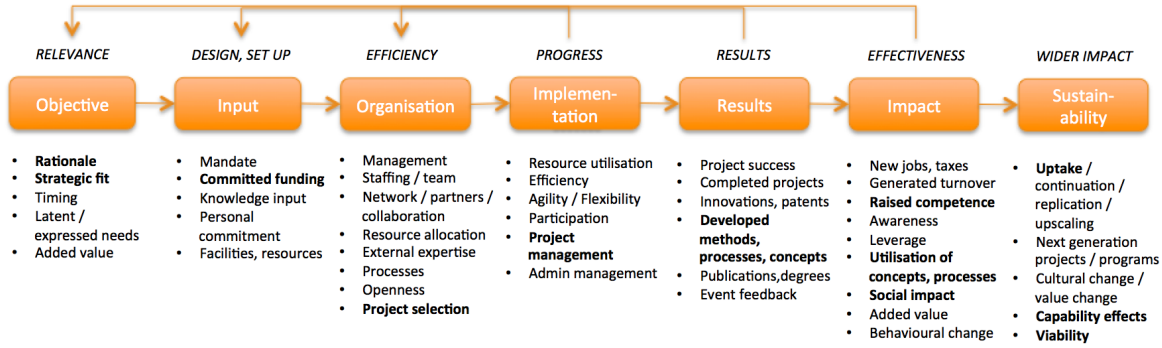

Figure 11. The evolution of typical monitoring and evaluation aspects over the life-cycle of a programme.

The evaluability analysis will include **a baseline analysis**, in which the reference values /conditions for monitoring, and for mid- and long-term impact management indicators will be defined. Some of the evaluation dimensions (such as impact on competence level) are likely to require that qualitative assessments are made of both the baseline level, as well as their progress, while as much as possible quantitative indicators will be utilised. See illustration below (Fig 12).

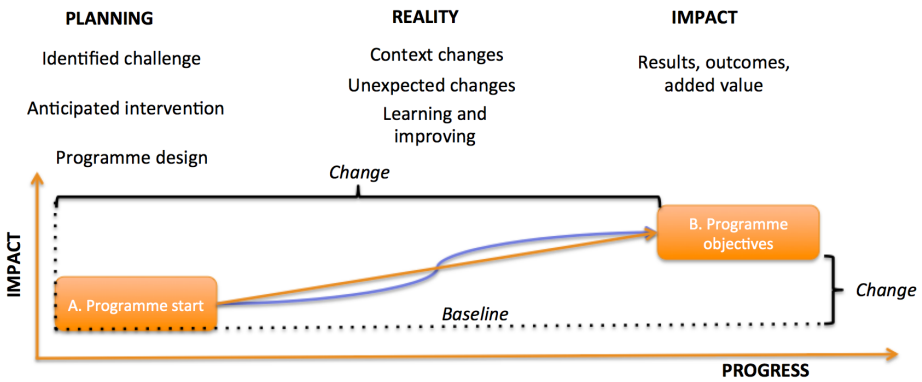

Figure 12. Illustration of the programme progress vs its impact and their baselines.

The extent of the baseline analysis will be discussed at the Evaluation Steering Committee, with regards to how extensively the baseline conditions are to be analysed in partner countries, and for the design and implementation of a **Baseline Survey**, which can be annually repeated (Impact Survey).

Evaluability analysis will also address the planned **collaboration practices** for the on-going evaluation and its utilisation in the programme management and will propose recommendations regarding those.

3.2.2 WP2: Meta-analysis and biannual reviews.

The second work package will start with a **meta-analysis** (2.1) of evaluation reports of the twelve MFA Innovation initiatives. Our team has a substantial insight regarding these initiatives and this will be now reflected in the systematic meta-analysis, drawing on the lessons, characteristics and overall trends of these initiatives.

Table 4. List of 12 programmes and the related documents to be covered in the meta-analysis.

Programme	Available reports and documents
1. The Energy and Environment Partnership Programme (AEA)	• *2012 Joint Mid-Term Review (In Spanish, Executive Summary in English)*
2. The Sustainable Forest Management Programme (MFS)	• *2012 Joint Mid-Term Review (In Spanish, Executive Summary in English)*
3. SAFIPA	• *2012 Evaluation memo* • *2011 Final evaluation report, 2 different versions of the same report* • *ToR*
4. BioFISA	• *2012 Mid-Term Evaluation* • *2011 ToR for Mid-Term Evaluation*
5. SAIS	• *2014 Mid-Term Evaluation* • *ToR MTE*
6. STIFIMO	• *2013 Mid-Term Evaluation* • *ToR*
7. Creating Sustainable Businesses in the Knowledge Economy (CSBKE)	• *2013 Mid-Term Evaluation*
8. Energy and Environment Partnership with Central America (“EEP")	• *2012 Mid-Term Review (In Spanish, Executive Summary in English)*
9. EEP S&EA Programme	• *2012 Joint Mid-Term Review*
10. EEP Mekong Programme	• *2012 Joint Mid-Term Review*
11. IPP Programme	• *2011 MTR*
12. TANZICT	• *2013 MTR* • *ToR MTR*

The key steps for the conduction of the meta-analysis are described in Table 5 below (tasks and resources are subject to adjustment).

Table 5. Description of the process, tasks and resources of the meta-analysis.

Task	Outcomes	Rough allocation of resources (days)
1. Specification of evaluation hypotheses	Elaboration of the anticipated outcome of the evaluation and the utilisation of results.	1
2. Design of the analytical framework and evaluation questions	An analytical framework that will describe what issues are to be looked from the evaluation reports and why, as well as what are the specific questions related to those issues, and how are the answers to be categorised.	3
3. Document analysis and initial reporting (powerpoint)	Systematic reading of the evaluation reports and documentation of the answers. To be discussed, whether complementary methods are used (e.g. interviews, additional documents).	10
4. Presentation of initial findings	Presentation of initial findings in a small workshop (e.g. extended steering group) to validate the results and to discuss implications, further analyses and reporting.	2

5. Reporting	Reporting of the key results in a concise form.	4

A major part of WP2 will be to conduct two **bi-annual reviews** (2.2) including desk analysis of the available documents and data. The analytical desk study of the review will foresee for the planning and conduction of respective field missions. The review mission plans will be proposed for the Steering Committee to approve.

A **standardised approach** and question structure / methodological approach will be utilised for each assignment, to ensure quality, efficiency and reasonable comparability of results. The key steps for the conduction of bi-annual reviews are described in Table 6 below (tasks and resources are subject to adjustment).

Table 6. Description of the process, tasks and resources of a bi-annual review (field mission).

Task	Outcomes	Rough allocation of resources (days)
1. Desk study and planning	Desk analysis of the available programme documents and data to form the basis of evaluation. Identification of key issues for review and methods of verification. Planning of the field mission.	3
2. Field mission	Mission arrangements, travel, conduction of meetings and interviews on site, documentation of outcomes and a short debriefing of the key findings of the mission. Conducted typically by two experts in parallel.	12
3. Reporting	Reporting of the review findings in a structured format (all reviews with a similar structure). One round of comments to draft report and finalisation of the report.	5

The mid-term evaluation of the BEAM programme (2.3) will consist of a summary of all the work carried out during the WP1 and WP2, with an emphasis on drawing on the overall conclusions of the programme progress so far. These findings and conclusions will be presented and discussed in a **workshop amongst BEAM Steering Group** and invited external experts. The outcomes and resources for the conduction of the mid-term review and workshop are described in Table 7 below (tasks and resources are subject to adjustment).

Table 7. Description of the process, tasks and resources of mid-term review and workshop.

Task	Outcomes	Rough allocation of resources (days)
1. Preparation of a draft mid-term review	The mid-term report will for a large part be compiled and synthesised from the various reports (bi-annual reviews, etc) produced by the evaluation team, reflecting the progress and performance against programme's initial targets.	10
2. Workshop with (extended) Steering Group	The draft mid-term review will be presented and the findings discussed at a workshop with an extended Steering Group. Feedback for finalising the report will be received.	2
3. Finalisation of the report	Finalising the mid-term report on the basis of received feedback.	3

It has been discussed in the Evaluation steering group that more evaluation resources could be allocated to the **on-going support** of the programme management. This has been noted, but such reallocations have not been proposed yet. A solution could be to replace one or

both of the bi-annual reviews in WP2 with on-going support activities , but this will be addressed later.

3.2.3 WP3: Biannual reviews in 2017-2019.

The option for WP3 will include additional five geographical and/or thematic biannual reviews (3.1-3.5) to be conducted according to similar specifications to WP2, and further to the decision by the programme Steering group. The reviews will be conducted all team members (5x20 days). All the seven reviews will eventually be synthesised into a Final Report (3.6), which summarises previous reports and draws on the lessons learnt and reflecting the changes made during the course of the programme, as well as analysing the generated impacts.

3.3 Evaluation schedule

Table 8. Schedule for working packages 1 and 2.

WP / Task	Wd	Oct 2015	Nov 2015	Dec 2015	Jan 2016	Feb 2016	Mar 2016	Apr 2016	May 2016	Jun 2016	Jul 2016	Aug 2016	Sep 2016	Oct 2016	Nov 2016	Dec 2016	Jan 2017	Feb 2017	Mar 2017	Apr 2017	May 2017	Jun 2017
WP 1: Evaluability of BEAM	**50**																					
1.1 State-of-the-art analysis	8	x																				
1.2 Analysis of ramp-up phase	8		x																			
1.3 Evaluability analysis	34		x	x	x																	
WP 2 : Meta-analysis&Reviews	**80**																					
2.1 Meta-analysis	20				x	x	x	x	x	x												
Planning	*3*				x																	
Analysis	*12*					x	x	x														
Reporting	*5*								x	x												
2.2 Biannual review 1	20											x	x	x								
Desk study & planning	*3*											x										
Field mission	*12*											x	x									
Reporting	*5*													x								
2.2 Biannual review 2	20																	x	x			
Desk study & planning	*3*																	x				
Field mission	*12*																	x	x			
Reporting	*5*																		x			
2.3 Workshop	5																			x		
2.3 Mid-term report	15																			x	x	x
TOTAL (WP1 & 2)	**130**																					

Table 9. Tentative schedule for working package 3.

WP / Task	Wd	Q3 / 2017	Q4 / 2017	Q1 / 2018	Q2 / 2018	Q3 / 2018	Q4 / 2018	Q1 / 2019	Q2 / 2019	Q3 / 2019	Q4 / 2019
WP 3 (optional, total)	**120**	x	x								
3.1 Biannual review 3	20	x									
Desk study & planning	*3*		x								
Field mission	*12*		x								
Reporting	*5*			x	x						
3.2 Biannual review 4	20			x							
Desk study & planning	*3*				x						
Field mission	*12*				x						
Reporting	*5*					x	x				
3.3 Biannual review 5	20					x					
Desk study & planning	*3*						x				
Field mission	*12*						x				
Reporting	*5*							x	x		
3.4 Biannual review 6	20							x			
Desk study & planning	*3*								x		
Field mission	*12*								x		
Reporting	*5*									x	x
3.5 Biannual review 7	20									x	
Desk study & planning	*3*										x
Field mission	*12*										x
Reporting	*5*										x
3.6 Final report	20										x
TOTAL (WP1, 2 & 3)	**250**										

Annex 1. List of references and extended literature

References

Albæk E. (2001), Vidensinteresser og de mange betydninger af evalueringet udviklingsperspektiv, in Dahler-Larsen P. & Krogstrup H. K. (red): Tendenser i evaluering. Odense: Odense Universitetsforlag. Translation in The Use of Evaluations in Europe Report from The European RTD Evaluation Network Meeting The Danish Institute for Studies in Research and Research Policy 2003/2.

Evaluative Learning Review Synthesis Report, USAID/CMM's People-to-People Reconciliation Fund, 2014.

Finns in the EU 6th Framework Programme, Evaluation of participation and networks, Tekes Reports 6/2008.

Gamble, J. a a. (2008). A Developmental Evaluation Primer. Report, 1–38.

Global Risks Report 2013, The World Economic Forum.

Hallie Preskill and Tanya Beer, Evaluating Social Innovation, 2012.

Lerner, J (2009) Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed--and What to Do About It. Princeton University Press, Princeton: NJ.

M. Scott Peck, The Road Less Traveled, 1978.

Michael Quinn Patton, Developmental Evaluation Applying Complexity Concepts to Enhance Innovation and Use, 2011

Michael Quinn Patton, Qualitative Research & Evaluation Methods, Integrating Theory and Practice, SAGE 2015

OECD. (2013). An International Benchmarking Analysis of Public Programmes for High-Growth Firms. OECD.

Quality Standards for Development Evaluation, OECD DAC, 2010.

Ray Pawson and Nicholas Tilley, Realistic Evaluation, SAGE 1997.

Ray Pawson, The Science of Evaluation. A Realist Manifesto, SAGE 2013

Results Based Management (RBM) in Finland's Development Cooperation – Concepts and Guiding Principles, MFA 2015.

Study on Network Analysis of the 7th Framework Programme Participation, European Commission, 2015.

Technopolis Group & Mioir. (2012). Evaluation of Innovation Activities: Guidance on methods and practices. Evaluation of Innovation Activities. Guidance on Methods and Practices. Study Funded by the European Commission, Directorate for Regional Policy., (June), 1–64.

Further literature

Behn, R. (2003). Why measure performance? Different purposes require different measures. Public Administration Review, 63(5), 586–606.
Cook, T. (2006). Collaborative action research within developmental evaluation. Evaluation 12(4): 418-436.

Horton et al (2003). Evaluating capacity development: Experiences from research and development organizations around the world. The Hague, Wageningen, Ottawa, ISNAR, CTA, IDRC.

Hunter, D. E. K., & Nielsen, S. B. (2013). Performance Management and Evaluation: Exploring Complementarities. New Directions for Evaluation, 2013(137), 7–17.

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.

Liket, K. C., Rey-Garcia, M., & Maas, K. E. H. (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.

Nagarajan, N., & Vanheukelen, M. (1997). Evaluating EU Expenditure Programmes. A Guide. Ex post and Intermediate evaluation.

McDavid, J., Huse, I., & Hawthorn, L. (2013). Program Evaluation and Performance Measurement: An Introduction to Practice. Sage Publications.

Patton, M.Q. (2012). Essentials of utilization focused evaluation. SAGE.

Patton, M.Q. (1994). Developmental evaluation. Evaluation practice 15(3), 311-319

Pawson, R., Wong, G., & Owen, L. (2005). Known Knowns, Known Unknowns, Unknown Unknowns The Predicament of Evidence-Based Policy. American Journal of Evaluation.

Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. Evaluation, 14(1), 29–48.

United Nations Development Group. (2011). Results-Based Management. Harmonizing RBM concepts and approaches for improved development results at country level. United Nations.

Van Dooren, W., Bouckaert, G., & Halligan, J. (2010). Performance management in the public sector. London & New York: Routledge.

Westley, F.; Zimmerman, B. and Patton, M.Q. (2006). Getting to maybe: How the world is changed. Toronto: Random House.