SVUF 2013


LITTLE PIECES OF A LARGE PUZZLE
Evaluation use and impact

Erik Jakobsson, APeL FoU
Mats Fred, Malmö högskola, CTA
## Contents

Evaluation – of what use? ................................................................. 3  
Aim of the paper............................................................................. 4  
Disposition.................................................................................... 4  
Ongoing evaluation ...................................................................... 4  
Evaluations of ERDF projects....................................................... 5  
Analytical perspectives ................................................................. 6  
Meta analysis.................................................................................. 8  
  Characteristics of the evaluations ............................................... 9  
  Active ownership...................................................................... 10  
  Case study | Syster Gudrun....................................................... 11  
Collaboration.................................................................................. 13  
  Case study | UMIT ................................................................. 13  
Developmental learning .............................................................. 15  
  Case study | FindIT ................................................................. 15  
Concluding analysis .................................................................... 17  
References.................................................................................... 20
Evaluation – of what use?

The interest in the practical significance of evaluation, evaluation use, has generated quite a lot of empirical research which also has generated several theoretical and conceptual categorizations (Mark & Henry 2004). Evert Vedung (1998 s.209f) describes six different concepts of evaluation use. The first, called the *engineering model*, treats the use of evaluation in an instrumental manner. Results from an evaluation are taken care of in a well-oiled machinery of decisions where social engineers quickly take action in accordance with the evaluation results. This model has been subjected to massive criticism and several researchers, including Vedung himself, say that it is not likely to occur to any significant extent. The *enlightenment model*, Vedung’s second concept, is a “softer” model in which the results of an evaluation not automatically lead to practical change, but rather acts as an informative instrument where politicians and other stakeholders obtain insights – problems are highlighted and the reality can be structured differently, using new concepts. A *legitimizing use* of evaluation is the third of Vedung’s six models. This means that politicians or decision makers use evaluations as a legitimizing means to justify decisions. Next Vedung describes an *interactive use* model in which users not only can get evaluation results but also help to create them. The evaluation is only part of a greater amount of insights that are important for the decision making. *Tactical use* is when an evaluation is assigned to give the impression that something is happening. It’s the assigning in itself of an evaluation that is the point, not that it results in something. The last type of evaluation use in Vedung’s list is *ritual use*, which means that we evaluate in order to be in accordance with modern thought.

Similar categorizations have been done by other researchers as well. Mark & Henry (2004) describe it in terms of *instrumental use, conceptual use, symbolic use, process use and misuse of evaluation*. Carol Weiss¹ (1980), an often cited researcher when it comes to evaluation use, describes six similar, but somewhat different categories: *Instrumental, knowledge driving, interactive, political, tactical and enlightenment* use of evaluation. Weiss categorizations come from a big empirical study conducted in the late 1970s where she investigated how decision makers react upon research and evaluation results. According to Weiss the researchers failed to locate any instrumental use of scientific knowledge in political decision making. They then drew the conclusion that the knowledge was irrelevant (Weiss 1980 .p396). However, Weiss pointed to the notion that the knowledge could in fact influence how the decision makers thought and by this have a more general and long term effect, what she called enlightenment.

The different categorizations describe different types of evaluation use. *Instrumental* and *enlightening* are theoretical ideal models of use whilst the *interactive* model for instance is a model of evaluation practice aiming to enhance the use of evaluation. Mobilization of support and *political use* deals with the rhetorical aspects of evaluation, or what Ove Karlsson (1999) describes as the declared intention of the evaluation in contrast to the actual use or function of the evaluation. All these categorizations deal with evaluation use but at different levels or from different perspectives.

¹ Weiss (1998) has a couple of years later summarized the six points in four main categories of evaluation use: *Guidance for action, reinforcement of prior beliefs, mobilization of support and enlightenment.*
Yet another way, and perhaps a simpler or possibly more practical way, to categorize evaluation use is done by Cousins (2003) who differs between process use and use of findings. The latter is about the result that comes out of an evaluation, often in the form of a report, which is presented for the client and then distributed through different channels. Process use, on the other hand, means that report in itself can be of less significance and the very evaluation process can constitute an opportunity to be used. This also means that even if the evaluation reports are not used as intended an evaluation can still be effective (cf. Forss, Reiben & Carlsson 2002, Vedung 2009).

**Aim of the paper**

In this paper we investigate an evaluation approach called ongoing evaluation, an approach advocated by the EU, and how this has been used during the Structural Funds programming period 2007–2013 in Sweden. Ongoing evaluation can be described as an approach to evaluation that promotes both evaluation use and evaluation results. We are interested in how this has been done in practice. Our empirical base for the analysis is a meta-analysis of 60 ERDF\(^2\) funded projects and their evaluations commissioned by the Swedish Agency for Economic and Regional Growth (Tillväxtverket). A theoretical framework, constructed originally to understand the sustainability of projects and program initiatives, is used in combination with program theory to investigate the use of ongoing evaluation.

**Disposition**

After this brief introduction we describe what ongoing evaluation is or is supposed to be and how it has been organized in Sweden. The section gives a context to the studied evaluations. After this we outline our analytical perspectives consisting of what sometimes is referred to as program theory along with a framework of three mechanisms for sustainable development work. In the following section we describe the meta-study of the evaluations and its results of the study. This also includes three illustrative case studies in connection to our theoretical framework. The paper ends with some concluding remarks.

**Ongoing evaluation**

The task of ongoing evaluation is to contribute to an effective implementation of the Structural Funds programmes in line with the revised Lisbon Strategy, the Europe 2020 Strategy and the Swedish national strategy for regional competitiveness, entrepreneurship and employment 2007–2013. This should be done by the evaluators identifying, documenting and communicating results that lead to long-term effects, structural changes and strategic impact. The Swedish national strategy emphasizes the importance of evaluation and learning:

---

\(^2\) European Regional Development Fund – ERDF – aims to strengthen economic and social cohesion in the European Union by correcting imbalances between its regions.
The Government’s intention is for a systematic follow-up and evaluation process to be part of the work carried out on the regional development policy and the European Cohesion Policy. The aim of this is to improve the program work from start to finish – from planning to implementation. There is a constant need to increase our knowledge of the world around us and of how different measures can best be combined in order to be effective and reach our goals. (The Ministry of Enterprise, Energy and Communications, 2007 p. 49)

The importance of evaluation and knowledge formation is further emphasized in Europe 2020, which is the EU’s growth strategy for the present decade. Europe 2020, consequently, stresses that in the next period the Structural Funds are to contribute to smart, sustainable and inclusive growth. This is dependent on the ability of so-called “smart specialization”, which means supporting a resource-efficient programme implementation and avoiding the duplicated funding of similar projects, a lack of synergies and inefficient ways of coordinating development efforts. This requires an extended knowledge formation among the different actors involved. Ongoing evaluation is thought of as one of the tools for these joint learning processes. The demand for a resource-effective implementation – in line with the ideas of smart specialization – calls for learning by ongoing evaluation at the regional level, between regions, in national arenas and at trans-regional level.

In the US so called Developmental Evaluation (Patton 2010) has become an appreciated concept. It’s an evaluation approach that supports innovation development to guide adaptation to emergent and dynamic realities in complex or uncertain environments and can assist social innovators develop social change initiatives. Innovations can take the form of new projects, programs, products, organizational changes, policy reforms, and system interventions. A complex system is characterized by a large number of interacting and interdependent elements in which there is no central control. This approach has some common ground with the idea of ongoing evaluation in the European context.

**Evaluations of ERDF projects**

At the regional programme level (the eight Swedish Regional Fund programmes), ongoing evaluation has shown that the programmes have made a difference to innovation (Brulin & Svensson 2012). New linkages and contacts have been established between academia, SMEs and large companies and a surprising amount of resources (given the Commission’s expectations on experimentation) have gone to R&D based innovation projects involving existing companies and industries. But the programme evaluations also point to an uncertainty for the future when it comes to financing innovation activities. For instance, in northern parts of Sweden innovation system projects are characterized by short-term financial solutions and high expectations on companies and organizations in the region when it comes to making use of the results (Tillväxtverket 2011a). The programme evaluations point to the risks involved when projects receive funding for three years but lack a clear idea of financing beyond this. In fact, this seems to be a fairly common situation. In a report from the Swedish Agency for Economic and Regional Growth (Tillväxtverket 2011b) some of the most common findings related to the ERDF projects are:

---

- The objective/target structure and project logic (theory of change) in many projects needs to be improved.
- Projects need to be defined more clearly in relation to regular activities and the overall objectives of the programmes.
- The division of mandates, roles and responsibilities in major projects needs to be clarified.
- A more active ownership of the projects is needed.
- Projects need to be better at developing learning structures in order to change, improve and strengthen the national regional growth policy.

The report states that more effective forms for the feedback of experiences and knowledge are necessary. More coordinated action for ongoing evaluation and learning in order to generate more benefits for regional growth is also requested. The important challenge is to create an evaluation culture that makes the implementation of results easier and which supports the initiation of new projects. Project owners need to be better at making demands and specifying requirements concerning the use of evaluation. The evaluators also need to be more professional in their role. Ongoing evaluation is a relatively new phenomenon, which can explain why the methods have to be improved (Svensson et al. eds. 2009).

In the preparations for the new programming period for the EU Structural Funds theory-based evaluation has come into focus more than before. Marielle Riché⁴ (2013) argues that during a programme’s life, theory-based evaluations are useful to assess if it progresses as planned towards outputs, intermediate and end outcomes, how it does it, and if it takes into account limiting factors in its implementation theory. Theory-based evaluations may analyse the results of experimental interventions launched to test new approaches. They also may examine the probability of programme effectiveness.

*Ex post, these evaluations can validate the assumptions underlying the programme theory against the actual outcomes by getting inside the “black box” and analysing the mechanisms leading to these outcomes (including unintended effects) and the specific context in which they work.* (Riché 2013 p. 66)

**Analytical perspectives**

Distinctions between activities, output, results and effects in development work or in a project can be identified in a logic model or programme theory that aims to clarify the logic behind a project (See e.g. Donaldson 2007, Funnell & Rogers 2011, Brulin & Svensson 2012). Which activities lead to which results and which results lead to which effects? Programme logic or programme theory is thus a structured way of working with causal relations in development efforts. This means that there is logic at work in the thinking and planning of a project in which the causal relations between the various elements and stages of a project are in focus. Every development effort brings about the first three steps of the program logic, but results and long term effects are more uncertain.

⁴ Marielle Riché is a senior evaluation officer at the European Commission, Directorate General for Regional and Urban Policy. She is responsible for commissioning and managing evaluations of the European Regional Development Fund and providing expertise to Managing Authorities on evaluation approaches and methods.
Brulin & Svensson (2012) argue that the most important conditions for achieving sustainable change can be summarized in three mechanisms, or driving forces, explaining causal relationships: active ownership, collaboration and developmental learning. The key element underpinning these mechanisms is ongoing feedback from experience and knowledge formation. These ideas have been presented in several reports and books, which have also formed the basis for training and courses in learning evaluation at several universities across Sweden. We have supplemented the model with what we call the institutional context, which we believe largely determines the conditions for the mechanisms for sustainability to be activated.

The first mechanism—active ownership—stresses the importance of an active owner concerned with, and with mandate to make use of, the results of a project. Several researchers (cf. Brulin & Svensson 2012, Forssell et al. 2013, Bakker et al. 2011) argue that a commitment from responsible decision makers is required in order to implement projects and for the logic of development to have an impact on the parent organization. The second mechanism, in the model above,—collaboration—focuses on the formation of joint knowledge. This is sprung from theories relating to innovation systems, networks and cluster formation (cf. Lundvall red. 1992, Porter 1998, Etzkowitz & Leydesdorff 1997, Laestadius, Nuur & Ylinenpää red. 2007, Laur, Klofsten & Bienkowska 2012). The main argument is that collaboration needs to be strategically organized and is not simply a matter of who signs what documents. The third mechanism—developmental learning—stands for learning that
promotes development, change and innovation and relates to organizational learning (cf. Senge 1990, March 1991, Elkjaer 2004, Ellström 2010). Developmental learning can create multiplier effects and when analyzing this mechanism theories of learning are combined with theories of implementation, dissemination and strategic impact.

Brulin and Svensson (2012) argue that development work leading to sustainable change can be studied as an interaction between these three mechanisms and we argue that the same mechanisms can be used as an analytical tool to investigate the use of evaluation. As already mentioned we have added institutional context to the model. This emanates from experience with evaluation work but also from contemporary research on projects and temporary organizations that stresses the need to emphasize the importance of the surroundings of organizations. This research argues that projects are not isolated islands (cf. Engwall 2003) secluded from their surroundings and that they are affected by and effecting their environments. There is a long tradition of research treating projects as isolated, technical, instrumental entities, easily transferable to all kinds of activities and organizational contexts as long as one has the right skills and employs the correct tools. The critics argue that projects are temporary organizations that need to be put in an organizational and societal context as they are a complex form of organization that is not easy dealt with through planning techniques. They should not be seen as unequivocal, demarcated assignments which can be pre-defined in terms of goals and means (cf. Sahlin-Andersson & Söderholm 2002, Hogdson & Cicmil 2008, Clegg et al. 2006).

Meta analysis

In this section we present the results of the study. The basis for this is a systematic review of evaluation reports from projects in Sweden together with three illustrative case studies of Swedish projects. The analysis aims to give a general picture of the evaluations. The case studies provide us with a deeper understanding of the role of the ongoing evaluations in the projects.

In all, the empirical base of the study includes 60 evaluations, which was equivalent to half of the project evaluations procured in ERDF projects in Sweden during the present programming period, at the time the study was made (2011-2012). The classification of the evaluation reports could be described as a meta-evaluation, a term introduced by Michael Scriven in 1969, and the need to audit evaluators work has since then increased (Pelli et al. 2008). Like evaluation meta-evaluation can be an ambiguous phenomena (cf. Vedung 2009). Statskontoret (The Swedish Agency for Public Management) (2001:22) describes three different kinds of meta-evaluations (cf. Stufflebeam & Shinkfield 2007): evaluation of an evaluation organization; evaluation of the quality and precision of one or more evaluations; summary of results from two or more evaluations. Our analysis has the character of the last one where the focus has been on what effect the evaluations have had on the projects.

The evaluation reports were all subjected to a systematic analysis (based on a qualitatively oriented internal questionnaire) from which we collected relevant information. Here it is important to note that the systematic review of reports does not tell the whole story about the projects or the evaluations. The ambition with the systematic review was to acquire a good overall picture of the
projects and the importance of the ongoing evaluations – how they were conducted and what impact they had on the projects.

As a complement to this we also conducted seven case studies in order to obtain a deeper understanding of the evaluations. The selection of cases was made on the basis of the review and our analytical framework. The case studies are largely based on document studies, although in several cases interviews with evaluators and/or and project managers were also conducted.

Characteristics of the evaluations

Major consulting firms are the most common category of evaluators, followed by small consultancies. Research groups in universities and colleges also occur as evaluators, but not very often. However, some university researchers are engaged by consulting firms. Many evaluators carry out their work using an interactive and consultative approach, but a significant proportion of the evaluations can be regarded as summative, with an emphasis on the end of the project period. This approach is contrary to the very idea of ongoing evaluation advocated by SAERG. The idea with ongoing evaluation is to follow the project from the beginning, work closely with the project (managers, steering groups and owners), give feedback from the evaluative processes and take part in the development of the project.

![Figure 3: The nature of evaluation](image)

Several evaluators have played a supporting role combined with a critical perspective (positioned in the middle, see Figure 3), although many have played a critically evaluative role but have not really supported the different processes of the projects (to the right in Figure 3). These evaluations could perhaps be compared with more traditional evaluations, where the evaluator takes a more distanced approach to the project and where the bulk of the work, and perhaps the most important part, connects directly to the final report. Quite a few evaluations are strongly supportive, but have no apparent critical, evaluative role. This role resembles that of an organizational consultant.

These differences in evaluation approaches are reflected in the analysis presented in the reports. More than half of the evaluation reports are quite descriptive in nature and do not contain much analysis. The most common approach, though, is reports based on an empirical analysis where conclusions are drawn from empirical observations. There are also examples of reports that actually use theoretical perspectives, concepts and models, although these are few and far between. In connection to this, or perhaps as a result of the strong descriptive and empirical nature of the

---

5 This also includes interviews and studies of documents.
reports, many evaluations focus on the activities, outcomes and short-term results of the projects, rather than on long-term effects and structural changes.

A lot of results are presented in the projects and the evaluation reports discuss them in different ways. The vast majority of evaluations answer the question of whether the project has achieved its goals or not. Many of the projects seem to have reached their goals and some have even exceeded them. We find descriptions of projects that have resulted in new ways of working, new forms of collaboration, new ways of disseminating knowledge and the development of new methods for change. However, there are also examples of projects that have had too many goals, where the logic (programme theory) has not fitted the project and where the results have mainly benefited universities.

With regard to long-term effects, we essentially found discussions about what the evaluator or representatives of the project believed the long-term effects would be, or how the project would lead to long-term effects. Some of the evaluators are unable to say much about long-term effects, since such effects can only be expected in the future and cannot be identified in the present. Some indicate that the project will continue as a solid business after the funding ends, while others point to less tangible things like processes of change that are emerging. There are also examples of evaluation reports that point to what needs to be done in order to achieve long-term effects. For example, some reports stress the need for collaboration between relevant players if the project is to lead to long-term effects, while others point to the need for the project owner to take responsibility for the results.

In general, the evaluations have a “local” character in terms of results and effects. Few of the evaluations deal with the overall objectives of the programme and the Swedish national strategy. The results and effects that are described are related to the immediate surroundings and the involved organizations.

When it comes to the horizontal criteria, these are usually given little or no space in the final evaluation reports, which indicates that the horizontal criteria are not considered to be of any real importance in the projects. It is somewhat remarkable that so many reports leave these criteria out altogether. In several cases the evaluation simply points to the notion that the criteria has no relevance in the project, or that the project simply does not deal with those issues. Very little criticism is directed towards the projects for not working with these criteria, and there is no discussion about what this might mean for the outcome of the project. This may have something to do with the vague connection of the horizontal criteria to the objectives of the programmes and the Swedish national strategy, although there are a few examples of some form of discussion and analysis concerning the horizontal criteria taking place.

Active ownership

Organizing an active ownership in large projects with a lot of actors seems to be difficult. An effective project organization does not appear out of the blue. In general, no in-depth discussion and analysis

---

6 Environment, gender equality and diversity.
of project organization, ownership and steering is found in the final evaluation reports from the projects. In fact, only a few of the reports we came across had any kind of critical discussion and analysis regarding this mechanism for sustainable change. In some cases there is absolutely nothing in the text about the project organization and how it works - not even at a very superficial level. We should point out that this does not necessarily mean that the projects do not have active owners or effective steering groups.

Many evaluation reports simply do not dwell on these issues. In fact, more than half of the final evaluation reports in the review do not address the issue of whether or not a steering group actually steers the project.

Some reports do reveal major shortcomings in terms of ownership and control, such as steering groups that do not actually control the overall focus and development of the projects, which might indicate a lack of active ownership. In many cases steering groups regard themselves as reference groups rather than forums with operational responsibilities. This seems to be a fairly common phenomenon. A weak point like this in the project organization can obviously have an important impact on the project’s ability to implement results and create long-term effects.

Below we present an illustrative case that demonstrate the importance of active ownership in order to maintain the sustainability of the changes and results that have occurred throughout the project.

Case study | Syster Gudrun

The project Syster Gudrun is a collaboration project between Blekinge County Council, Blekinge Institute of Technology, Affärsverken i Karlskrona\(^7\) and the Municipality of Karlskrona aimed at creating a full-scale information technology labs in health care organizations. The basic idea has been to find out whether and how modern technology could make health care more accessible without a loss of quality for the individuals concerned, and at the same increase efficiency. The short-term aims of the project are:

- To create production-oriented full-scale information technology labs in health care organizations in Blekinge and parts of Skåne.
- To build a full-scale research-based lab at the Blekinge institute of Technology.
- Develop a web-based platform for dialogue regarding public health care as a complement to the national telephone-based health care advisory service.
- Increase IT competence health care and thereby serve as inspiration, a disseminator of ideas and an educator.
- Professional development of health care employees.

---

\(^7\) Affärsverken is a Karlskrona-based company that builds, develops and operates infrastructure in the municipality of Karlskrona (www.affarsverken.se/english).
According to the evaluators the project has achieved these goals. The production-oriented and the research based labs are up and running and will continue to operate after the project funding ends. The web-based platform for dialogue has been developed and the receiving organizations are ready to take over. The use of IT in rehabilitation and neonatal care has been implemented at Blekinge Hospital. The planning of health care at a distance has also been implemented at Blekinge Hospital and in several municipalities in Blekinge.

The project has also, according to the evaluators, managed to generate considerable interest in the different media\(^8\) and the evaluators themselves have contributed to a public debate by publishing articles.

The evaluators interacted closely with the project’s management and owners throughout the project process. They gave interim reports every six months in which the project and its development were described. The interim reports also had different themes. For instance, one of the reports focused on the processing role of the evaluation, another on the steering group and a third on the horizontal criteria. All the reports, including the final report, describe the advantages and disadvantages of the project and include suggestions about how to move forward. The evaluators also took part in meetings with the steering group and conducted nine seminars with the project staff and the steering group. The seminars seem to have had a huge influence on the project. Based on interviews, observations and/or research studies, the seminars provided an arena for learning where discussions could be held on the basis of empirical or theoretical perspectives and serve as a basis for change. This joint learning led to the evaluation becoming an important tool for development in the project and has in fact changed the project’s organization and its way of working.

One major impact of the evaluator’s work concerns the functioning of the steering group and the ownership of the project. This also includes discussions about the implementation of the results. Early on in the process the evaluators noted that the project had a dysfunctional steering group. By this the evaluators meant that the steering group did not consist of people with a mandate to engage in strategic discussions and decision-making at the organizational level. There were initially in fact no obvious recipients of the results and the steering group lacked representation from organizations with a mandate to make decisions.

The evaluators pointed to the importance of a steering group represented by actors with an interest in the project and the ability to make actual changes in the recipient organizations. This was also the focus of one of the interim reports and one of the seminars. These discussions had an impact on the project and the composition of the steering group was changed. This change, the evaluators believe, had an impact on the overall results of the project and made it easier to implement the overall project results, as well as those from the sub-projects, in the recipient organizations.

\(^8\) [http://svt.se/svt/jsp/Crosslink.jsp?d=33782&z=839402](http://svt.se/svt/jsp/Crosslink.jsp?d=33782&z=839402)
[http://www.ltblekinge.se/omlandstinget/mediaservice/pressmeddelanden/pressmeddelanden2011/systergudrunakertillalmedalen.5.39634a231309b646bccc8000544.html](http://www.ltblekinge.se/omlandstinget/mediaservice/pressmeddelanden/pressmeddelanden2011/systergudrunakertillalmedalen.5.39634a231309b646bccc8000544.html)
The evaluators also experienced that a simultaneous combination of the role of critic and supporter was not an easy task, especially when suggesting major changes in the organization of the project. The building up of trust requires a certain amount of closeness to the project, although distance is also necessary in order to criticize when this is necessary.

Another example of the ongoing evaluation’s impact on the project relates to the horizontal criteria. Early on in the project the evaluators noticed that no efforts were being made in relation to the horizontal criteria and asked why this was the case, both in an interim report and in a seminar. As in the case of the steering group, this changed the project’s way of working. Guidelines were developed for how to work with these issues and make them visible. It is therefore clear that the ongoing evaluation provided the project with an arena for discussion and development.

**Collaboration**

Regarding the descriptions and analysis of collaboration in the evaluation reports there is more to be desired. As mentioned above, the issue of project organization in many evaluation reports is seldom discussed. The most common type of collaboration in the projects themselves and in the innovation environment in which the projects take place can be characterized as Triple Helix, i.e. collaboration between academia, industry and public agencies. In most cases the collaboration seems to be formalized and binding and more often than not the dominant actors in the Triple Helixes are the public sector and/or academia. In several projects the concept of Triple Helix has little or no relevance, for the simple reason that one of the spheres of the Triple Helix is missing, usually academia or industry.

Collaboration seems to be an important mechanism for sustainable change. However, organizing a dynamic, effective and innovative collaboration between different stakeholders is a difficult task. In the following case we take a closer look at how a project dealt with these issues.

**Case study | UMIT**

UMIT⁹, a strategic project, was initiated in 2009 by Umeå University within the focus area Applied Information Technology. UMIT Research Laboratory is an environment for research in computational science and engineering with a focus on industrial applications and simulation technology. The project has facilitated the bringing together of researchers in a new research laboratory at Umeå University, which opened in May 2011. The vision of the UMIT Research Lab is to become a world-leading platform for software development and attract companies like ABB, Ericsson, Google and Microsoft.

The UMIT Research Lab focuses on interdisciplinary research, education and collaboration with industry and society. The research results create new opportunities for technological and scientific simulations and completely new software and services for the processing of large amounts of information. Deliverables include new models, methods, parallel algorithms and high-quality software targeting emerging HPC platforms and IT infrastructures. Solutions are tested together with

---

⁹ UMIT stands for Umeå University, Mathematics, Information and Technology.
industry partners in sharp application projects targeted at new competitive products and job opportunities.

Funders of UMIT are The Swedish Research Council, the Baltic Foundation, Umeå Municipality, ERDF and Umeå University. ProcessIT Innovations\textsuperscript{10} and industry companies co-finance specific project activities. Examples of industries in which UMIT has partners are mining, paper, automotive and IT. Currently, UMIT includes seven research groups with 45 researchers and project staff. Six professors are funded by The Swedish Research Council. UMIT operates 20 projects with more than 20 industry partners. One ambition is to have a staff of 75 people by 2015.

After a little more than two years UMIT has achieved significant results. Obviously the research laboratory in itself is one. Furthermore, a trainee programme – UMIT Industrial R&D Trainee Programme – has been started. The project has conducted 12 recruitments to the laboratory, including a research manager, a research group with PhD students and post docs and two associate professors. Also, the UMIT activities have led to the creation of about 20 external jobs in industry and academia. The project has also generated new techniques and software. In addition, UMIT has added value to industry partners e.g. in terms of new business opportunities, assistance in product development, energy savings and enhanced competitiveness.

The idea is that UMIT will become a permanent unit within the university and, furthermore, become a world leader in its field. This will be achieved by developing strong and vibrant collaborations between key players in academia, industry and the region. The evaluators’ efforts helped UMIT see that the collaborations needed to be strengthened and the evaluators therefore suggested that they should conduct an analysis of successful environments like UMIT in order to learn more about collaboration and what is necessary for long-term sustainability. The evaluators thus explored environments that have been successful in developing collaborations and creating strong organizations by interviewing strategically selected individuals representing ministries, agencies, funders and innovation environments. The evaluators also contributed with more theoretical references and perspectives, based on the literature in this field. This presents an interesting example of how ongoing evaluation can be instrumental in strengthening a key mechanism in development work, namely collaboration.

The evaluators’ conclusions concerning strategic collaboration were as follows. Successful collaboration environments:

- are clearly characterized by the university’s history, culture and role in the specific region
- are organized communities characterized by mutual understanding and mobilization of involved parties
- conduct strategic outreach and engage in strategic communication
- provide a strong culture characterized by entrepreneurial incentives and formalized partnerships.

On this basis the evaluators made the following recommendations to the project:

\textsuperscript{10} ProcessIT Innovations is a collaboration center at Luleå University of Technology.
- Develop long-term strategies and strive to become an integral part of the regional agenda.
- Create mutual mobilization by becoming a natural meeting place for all involved.
- Work strategically with outreach activities and develop strategic communication.
- Strengthen the UMIT-culture to minimize dependence on individuals and take advantage of the entrepreneurial spirit that characterizes the environment.

The study and the recommendations have been perceived as valuable and useful by the project management. On the whole, it would seem that the interventions of the evaluators have been appreciated and deemed valuable by the project management. Before undertaking the strategic task described above, the evaluators provided the project with an analysis of the current situation. They have also helped the project to focus on the long-term goals, even beyond the project period, and to focus on its performance indicators. The evaluators also conducted follow-ups by interviewing many of the actors in UMIT’s networks, which among other things showed that the support for UMIT was great but that UMIT’s role in the regional innovation system was somewhat unclear.

**Developmental learning**

Very few final reports describe learning processes or activities that are specifically aimed at learning, and few describe ongoing evaluation as a mechanism for change in the projects. Despite this, it is reasonable to assume that different kinds of learning and change have taken place as a result of the evaluators’ activities and reports. Many evaluators are not outspoken or explicit about this in their reports, but several do mention the importance of ongoing dialogue with the project leaders. Several evaluators have also produced interim reports. The concept of learning rarely appears in the evaluation reports. Almost no reports at all discuss the multiplier effects of the projects and there seems to be a lack of awareness that ERDF programmes, according to the national strategy, are supposed to promote learning that leads to multiplier effects. However, in some cases it would appear that processes leading to multiplier effects have taken place, even though the concept of multiplier effects is not actually used.

As already shown in the above examples, learning through ongoing evaluation can be instrumental in improving the projects. But innovative change cannot be organized in a linear way. There must be room for experimentation, chance and adaptation to new circumstances. Below we examine a project where developmental learning has been an important factor, both for the results produced and for the sustainability of the project.

**Case study | FindIT**

The project was an effort to further develop a competence centre in the field of industrial IT based on the needs of small and medium-sized businesses in the Gävleborg region. The overall goal was to strengthen competitiveness through smarter and more efficient business system solutions. Cooperation between companies and industries on industrial IT was an important element of the
project. The project owner was Sandviken Municipality, through the technology park, Sandbacka Park.

FindIT spans several different industries and target groups. This has been a difficult challenge for the project because all these different categories have different information needs and development interests. It has required great efforts on the part of the project to identify and understand the various categories and what their interests and needs are within the broad field of industrial IT. The evaluators believe that the project has now identified the relevant categories. Contacts with the small and medium-sized enterprises have been expanded both quantitatively and through enhanced information exchange. The evaluators perceive the enhanced exchange of information as an indication that the companies’ confidence in and expectations of the project have increased. In other words, the project’s legitimacy to act in the business has been strengthened.

FindIT illustrates how developmental learning can be achieved through ongoing evaluation. The evaluators have in fact contributed to a new understanding of what is important and worthwhile for the future, even if this goes against the logic of the regional fund programme. The project has, with the help of the evaluators, come to realize that it is important to find complementary ways of evaluating the results and impact of the project. This means highlighting issues of efficiency and the long-term competitiveness of the SMEs, instead of focusing primarily on indicators in terms of new jobs and new businesses. The evaluators had a crucial role to play in illustrating the effects of the project and also those that may occur after the project period is complete.

Firstly, this new understanding means that the number of employees working with business system solutions needs to be reduced, and that the competence of the employees needs to be increased. Secondly, the 170 companies working with business systems solutions in the region probably need to be reduced by half and significantly upgraded in terms of competence in order to be able to collaborate with international suppliers such as SAP. Thirdly, FindIT needs to make sure that business system solutions, such as innovation in services, become a natural feature among the companies in the region. When e.g. the headquarters of the large engineering group Sandvik moves to Stockholm the challenge will be to upgrade in the Gävleborg region in order to show that the region is at the forefront of service innovations, such as business systems development. This may be the only way of working towards a more competitive region.

The evaluators made the project management aware of the fact that, in order to be successful, FindIT probably needed to actively worsen the core indicators: new jobs and new firms! However at a systems level, FindIT has delivered results in accordance with the overall quantitative targets specified in the project application. However, in the longer term it is likely that there will be fewer companies and fewer employees, but better skills.

---

SAP is the world’s leading provider of business software.
Concluding analysis

We have, in this paper, showed that ongoing evaluations have had an impact on the development of several projects. However, it is also evident that the role of the evaluators is very diverse in its character. It varies from a traditional, distant, critical, summative, evaluating role to one of an organizational consultant working more as a “critical friend” or sometimes even as an assistant manager. In the middle of that range we found some examples of evaluations with a balance between closeness and distance, critique and support and evaluations that indeed did have an impact on the projects and their ability to lead to long-term effects.

As we have already touched upon we believe that the mechanism named developmental learning stresses the importance of the organizational structure and culture in which the projects are situated. We did find projects working with organizational learning, and we believe such projects fit well with the idea of ongoing evaluation and can make more use of it. Hertting and Vedung (2009) make a similar remark and argue that evaluations should be managed in accordance with the steering doctrine of the organization. In short the argument goes like this: If an organization is characterized by performance management the evaluation should be organized with a focus on performance. In line with that argument ongoing evaluation is best suited for organizations characterized by organizational learning and organizational development. By recommending a certain evaluation approach such as ongoing evaluation (in the form that it has taken in Sweden) you also to some extent indirectly recommend a certain way of organizing projects. However, in this paper we argue that one should pay attention not only to the organizational context but also to the institutional context. The institutional context can have prohibitive as well as enabling effects on the mechanisms for sustainability advocated by Brulin and Svensson (2012).

In the strategic discussions for the upcoming Structural Fund programming period (2014-2020) even greater emphasis has been put on theory-based evaluation:

Theory-based evaluations can provide a precious and rare commodity, insights into why things work, or don’t and under what circumstances. The main focus is not a counterfactual (“how things would have been without”) rather a theory of change (“did things work as expected to produce the desired change”). The centrality of the theory of change justifies calling this approach theory-based impact evaluation. (European Commission 2013 p. 7)

The accentuation of theory comes, we believe, from the fact that the evaluation practise during 2007-2013 to a large extent has been non-theoretical. This is in line with the research presented in this paper. We did find theory-driven approaches but the examples were quite few. In the case of UMIT we saw how the evaluation contributed with theoretical perspectives, based on relevant literature in the field, as one example. However, the majority of the reports in our study had a descriptive and/or empirical focus to their analyses, which tends to make the results very case-specific and therefore difficult to draw more general conclusions from. The problem with this is, according to Rothstien (2002), that this research or evaluation often lacks the ability to say something that the practitioners are not already well aware of.
Exactly how the theory based evaluation approach will end up looking we, of course, don’t know. Possibly it will enhance the use of research or theory to understand and help projects and programs work better, and thereby achieving the objectives stated. But it could also indicate a use of more scientific methods in conducting evaluations. As a combination of these two, Donaldson (2007) argue for what he calls program theory-driven evaluation science:

Program Theory-Driven Evaluation science is the systematic use of substantive knowledge about the phenomena under investigation and scientific methods to improve, to produce knowledge and feedback about, and to determine the merit, worth, and significance of evaluands such as social, educational, health, community, and organizational programs (Donaldson 2007 p. 10).

Donaldson (2007) uses a simple three-step model for understanding the basic activities of program theory-driven evaluation science:

1. Developing program impact theory.
2. Formulating and prioritizing evaluation questions.
3. Answering evaluation questions.

The phrase program theory-driven (instead of theory-driven) is intended to clarify the meaning of the use of the word “theory” in this context and aspires to clarify the type of theory that is guiding the evaluation questions and design. Evaluation science (instead of evaluation) is intended to underscore the use of rigorous scientific methods. Program theory-driven evaluation science is essentially method neutral. The focus on the development of program theory and evaluation questions frees the evaluators from method constraints. Program theory-driven evaluation science forces evaluators to try to understand the evaluand (the program) before rushing into action. The evaluator becomes familiar with the substantive domains related to the problems of interest in order to help formulate, prioritize, and answer key evaluation questions. This way one fundamental challenge in evaluation practice can be addressed – that is the risk of wrongfully reporting null results from a program, due to lack of understanding of and sensitivity to the program and the context (Donaldson 2007).

As we understand it a theory-driven evaluation approach does not mean that the concept of ongoing evaluation will be abolished. Rather it will be developed further. Ongoing evaluation has made an impact during the programming period 2007-2013. The case studies we conducted (of which three have been presented here) illustrate how ongoing evaluations have brought about changes in some projects. These changes have been of great importance in terms of the projects being able to generate long-term effects. The most obvious problems that the evaluators have detected and analyzed are listed below. The problems are largely interrelated.

- Deficiencies concerning organization and steering have prevented projects from generating results and effects that were potentially achievable. Through the efforts of ongoing evaluators the organization and steering of projects have been improved, which has resulted in projects performing better.
- Too many or too divergent ambitions and goals and/or too much focus on activities and short-term results have contributed to projects losing sight of their overall longterm goals. Ongoing evaluations have helped projects to sort out what is really important and what
should be prioritized in order to achieve the long-term targets set, even if this to some extent goes against the quantitatively-oriented logic that is included in programme and project descriptions.

- Effective collaboration inspired by the concept of Triple Helix is difficult to achieve. Projects have struggled with this task in different ways. Ongoing evaluations have contributed with important knowledge input and new insights to projects in order to make collaboration efforts more successful.

- Projects inspired by the concept of Triple Helix have put little emphasis on the later stages of innovation, e.g. developing new business models, although this is what participating entrepreneurs often need. Ongoing evaluations have pointed to the need for more demand-driven innovation efforts.

In sum we argue that ongoing evaluation has demonstrated merit, but our meta-analysis has also indicated some shortcomings associated with the evaluation practice during the current programming period. There is significant room for improvement and not least a need for the evaluations to aggregate more general knowledge from development efforts such as those described in this paper. The emphasis on theory-based evaluation for the next programming period of the EU Structural Funds is promising in that respect.
References


