PERGAMON

Evaluation and Program Planning 22 (1999) 73–82

EVALUATION and PROGRAM PLANNING

Using program theory models in evaluation of industrial modernization programs: three case studies

Hans Torvatn*

SINTEF IFIM, 7034 Trondheim, Norway

Accepted 1 August 1998

Abstract

Program theory oriented evaluations involve construction of a model that describes the logic and context of the program and enables the evaluator to check on program progress and impact before the program is conducted. Such constructions could be done in a multitude of ways: this paper describes the use of a particular tool called chains of reasoning. Chains of reasoning is a combination of text and a graphic image which presents the activities of a program or project, the goals to be attained, additional assumptions and the links between them. The paper uses three cases studies from Norwegian industrial modernization programs to demonstrate the strength of this tool in different settings. The first case study is a formative evaluation, the second an impact/summative evaluation and the third an additionality analysis. The paper concludes with a discussion of the strengths and weaknesses of the tool. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Program theory; Chains of reasoning; Industrial modernization evaluation

1. Introduction

The aim of this paper is to provide examples of program theory models use in economic and technology development programs. The paper will provide a brief introduction to program theory models, and the rationale behind such evaluation models. It will also present a particular tool for program theory evaluation called chains of reasoning.

The paper will then describe use of this particular tool in three cases of Norwegian industrial modernization programs. One case describes chains of reasoning used as a tool for program development in an early phase, a second describes chains of reasoning in impact evaluation, and a third case will describe how to use chains of reasoning in additionality analysis. Finally some recommendations on where to use and where not to use chains of reasoning are presented.

2. Program theory models and chains of reasoning

Program theory is an evaluation model originating from Huey-tsyh Chen and Peter Rossi (Chen, 1990; Chen & Rossi, 1983, 1987, 1992) and their works in the 1980s and early 90s. Over the years, the model have become quite popular all over the world, and various variants have been developed (Weiss, 1996).

According to Chen, a program theory is "a specification of what must be done to achieve the desired goals, and what other important aspects may also be anticipated and how these goals and impacts may be generated" (Chen, 1990, p. 43). Program theory should include both descriptive and prescriptive elements. The prescriptive (or normative) part of the theory should prescribe what should be the essential components of the program, how the program should be implemented and what goals should be pursued. On the descriptive side, program theory explains the causal processes underlying the programs. A key element in program theory is the idea that theory based evaluations could be used to strengthen the validity of evaluation without random assignment (Weiss, 1996). This is of particular interest in evaluation of industrial modernization efforts. In such efforts self-selection of participants make randomized experiments impossible, and while some quasi experiments have been carried out they remain rare.

The theory driven evaluation should use the program theory to identify central issues in the program, select research methods for data collection and data analysis, as opposed to a method driven evaluation where research

* Tel.: +47-73-596127; fax: +47-73-2570; e-mail: hans.torvatn@ifim.sintef.no

PII: S0149-7189(98)00043-3

steps are guided by the application of the particular method chosen.

In short program theory is a model that describes the logic and context of the program and enables the evaluator to check on program progress and impact before the program is conducted. A program theory driven evaluation is one where the evaluator constructs a program theory and uses this theory as a guide in the evaluation process.

The construction process could be done in a multitude of ways. Various models have been developed to aid in this construction process (Rush & Ogborne, 1991). One of them, called chains of reasoning, will be described in detail in this paper.

According to Finne, Levin and Nilssen (1995, p. 23) a chain of reasoning is: "a combination of text and a graphic image which presents the activities of a program or project, the goals to be attained, additional assumptions and the links between them. Each actor may have a different chain of reasoning, or program theory, or cognitive map, which integrates their thoughts of the program". It is dynamic, and as time goes, it will also include unintended impacts. A chain of reasoning is only one element in a complete program theory. A complete program theory also encompasses elements not included in chains of reasoning such as a description of program context, the rationale underlying the program, experiences from similar earlier programs and propositions about the program, if any. However, we will concentrate on chains of reasoning in this paper, as the most central tool in the construction process of a program theory.

3. The context: Norwegian programs on business development and technology transfer

As for most other industrialized countries the Norwegian government employs a substantial set of policies and programs to assist the Norwegian industry, and in particular the manufacturing industry, to assist them to adopt and deploy enhanced technologies and practices. In general it seems that Norwegian policies on industrial modernization are similar to U.S. policies in this area as described by (Shapira, Youtie & Rossner, 1996), and the politics of other European countries as described by (Georghiou, 1995; Rothwell & Dodgson, 1992). There are of course some important differences, due to structural differences among the various industries, governmental levels and prominence of these politics. At the same time, policies originating in Denmark have spread to both U.S. (Rosenfeld, 1996) and Norway, and Norwegian efforts have been 'exported' to other parts of Europe, thus demonstrating that the problems to be solved and solutions to solve them are indeed similar. Regardless of the policies in question and the country where the policy is implemented, the overall goal remains the same: improved competitiveness and economic performance. Most of the efforts are directed towards small and medium sized enterprises (SME), which in Norway usually are defined as firms with 100 or fewer employees.

Following the taxonomy introduced by Georghiou (1998) the efforts can be divided in two main categories: financial support and opportunity enhancing innovation policies. The first category covers direct and indirect economic support. Among these are all kinds of subsidies, risk reducing efforts, low interest loans, investment supports, tax reduction for specific activities and so on. The other category covers policies which create or enhance opportunities for innovation. Among these are institutions supporting innovation and technology transfer measures. Technology transfer need not be limited to technological and engineering knowledge, as will be demonstrated in this paper, management skills are also part of technology transfer. One of the main trends in European innovation policies during the 90s has been a move away from financial support towards opportunity enhancing efforts (Georghiou, 1998). This trend has been noted in Norway as well (Arbo, 1993).

In Norway there is no lack of opportunity enhancing efforts and has not been for the last decade. A survey by Torvatn and Munkeby (1994) identified 24 large industrial modernization programs. The total annual public funding of these 24 programs amounted to 70 million dollars. In addition to these 24 programs there are numerous smaller programs, local and regional programs (not included in the survey), programs operated by the department of agriculture[1] and programs stimulating export. A less comprehensive overview in 1997 (Torvatn, 1997) indicated that although individual programs come and go, the total volume of programs and expenditure on such efforts were still increasing. Since all these efforts are targeted at a population of some 7–10,000 firms, the program environment for industrial modernization programs in Norway is rather competitive. Recruiting participants is rarely easy.

Although not mandatory by any law or regulation there is a strong tradition of evaluation of industrial modernization programs in Norway. All the programs identified in the various surveys have been evaluated. The use of chains of reasoning in three of these evaluations will be presented as case studies.

4. Program A: chains of reasoning as a tool for program development

4.1. Description of Program A

Program A[2] was a product development program, targeted specifically at the food and food processing indus-

[1] For the benefit of the food industry.

[2] Official Norwegian name of Program A: Produktutviklingsnettverk (PU-nettverk). In English: Product development networks.

try. The goal of the program was to stimulate innovation among participants by teaching them product development methods. The program was operated by the Norwegian Institute of Food Research, and focused on food technical methods concerning the production process and the quality of the products. The program consisted of a series of seminars where product development methods were presented to the participants. Between each seminar the participants should do some homework and test out the methods on some product they were developing. To assist them in this homework each participant could claim assistance from a mentor from The Institute of Food Research, and were instructed to cooperate with the other participants.

The evaluation of the program was carried out in two steps. First a continuous formative evaluation of a pilot project of the program, to help develop a robust program (Torvatn, 1996). Then followed an impact oriented evaluation one year later (Torvatn & Elvemo, 1997). In both evaluation projects a program theory was constructed using chains of reasoning, however in this paper only the first formative evaluation will be presented as an example of how chains of reasoning could be used in program development.

4.2. *Use of chains of reasoning in Program A*

From the outset the program management had rather strong and clear models of how the programs activities and goals were linked. Together with the evaluator the following chain of reasoning was constructed for Program A:

The chain of reasoning in Fig. 1 should be read as follows: filled (black) arrows show the programs primary logic, in other words the 'main roads to success'. As can be seen from Fig. 1, Program A had two important ways of goal attainment. The first was through the seminars, where the participants should learn new product development methods and employ them in their product development process, together with subsidized product development services (hopefully from the Institute of Food Research) and thus improve their product development. Parallel to this the participants were expected to form a network among themselves and exchange experiences on their product development work, thus adding to their learning. Also note that the program expected to be able to choose between possible participants, and had developed criteria for selection.

The open (white) arrows show secondary/less important roads to success. Program A appointed a mentor to each participant, the mentor was expected to assist the participants if necessary. Also it was expected that the participants would learn something about organization of product development through the seminar and exchange of experiences, and this would hopefully also improve their product development work.

Regarding the vision of improved competitiveness the program operator (as well as the funding agencies) had a long term perspective on this, and did not expect any immediate results in that area.

Having constructed this chain of reasoning the program and the evaluator started working. The program began to recruit and carried out its pilot set of seminars. The evaluator trailed the pilot program, as participant observer. When the pilot project was completed the evaluator constructed a new chain of reasoning for the program, this one to sum up the learning for the program. This chain of reasoning is presented below:

In Fig. 2 the evaluator sums up the work of the pilot project of Program A in a chain of reasoning. As can be seen by comparing Figs 1 and 2, the program had not worked exactly as planned.

For one thing the program had not been able to attract enough companies in order to employ its selection criteria. The only real selection criteria had been willingness to participate. This was not surprising given the competition among industrial modernization programs in Norway.

Neither had the program been able to create any network between the participant, except for discussions during the seminars. The program had been able to teach the companies new product development methods. The methods had been employed by the companies. However, the companies felt they could do with more assistance from the mentor, both in employing new methods and organizing their product development work. (Hence the open arrows from the mentor.) In general the participants asked for much more assistance by the mentor, thus making it more important to select mentors. The program had not had clear criteria for selecting mentors, because the function originally had been judged unimportant.

The participants had not bought any subsidized product services. They claimed that this was too early for them, but they wanted the opportunity and asked the program to extend the period in which the offer of subsidized services was eligible. The evaluation was completed at a point in time when it was a little bit too early to tell whether or not the goals had been attained, but the program had had some progress which in due time could lead to the desired results (as expressed in the open arrows).

Based on this program and the evaluator agreed on the following changes in the program. First of all the mentor function must be strengthened. The program instituted training and developed new criteria for selecting mentors. Further, the networking between the companies had not functioned as planned, the companies being competitors seeing no reason to create a joint product development network. However, they were willing to exchange experiences from each other during the seminars, thus the program could emphasize this type of learning in future programs.

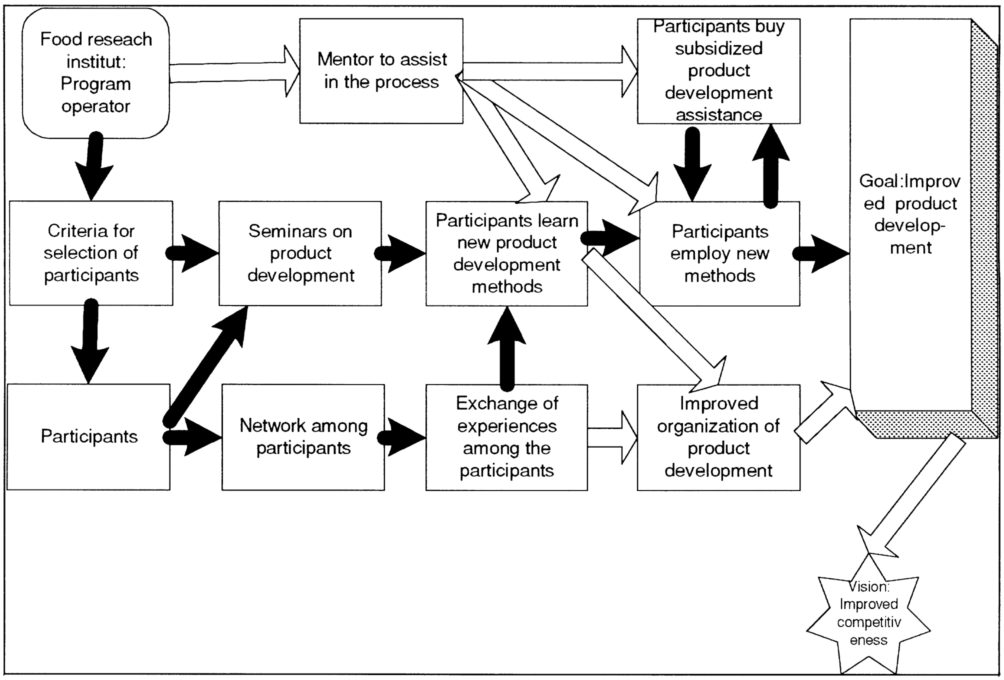

Fig. 1. Original chain of reasoning, Program A (from Torvatn, 1996).

In this evaluation chains of reasoning functioned as a tool for clarifying ideas, for summarizing evaluation results and as a vehicle for communicating the same results. The evaluation used the first chain of reasoning (Fig. 1) mainly to clarify the underlying models guiding the program management, and identify key issues in the program. By constructing a revised chain of reasoning as part of the evaluation it was easy for the evaluator to track progress, identify problem areas and suggest improvements. More important, the chains of reasoning created a basis for discussions about the program, which the program management could understand. The recommendations in the evaluation were easily communicated to, and accepted by the program management.

5. Program B[3]: Chains of reasoning in summative evaluation

5.1. Description of Program B

Program B was a four-year 12 million dollar effort by the Research Council of Norway to 'close the technology gap' between the available technology in the research institutes and the SMEs. Program B was to stimulate and increase the use of new technology by identifying new technology that would fit into the strategic plan of the companies, thus reversing the old model of finding companies that could use a specific technology. Since few of the participating SMEs had any strategic plan, a core activity of Program B became to assist the firms in a strategic development process. Having a strategic plan, the companies should identify and prioritize between a set of measures to undertake, in order to improve competitiveness. This process should lead to a 'Plan of action', which completed Program B in the company. Implementation of the plan was left to the companies.

This process was carried out by consultants trained and screened by the program. A total of 312 assignments were carried out. The program was evaluated continuously during its lifetime, and was generally judged successful (Finne, Levin & Nilssen, 1993, 1995). In fact the concept of Program B (training small business in strategic development) was 'exported' to several business development programs in eastern and western Europe.

5.2. Use of chains of reasoning in Program B

Before the impact evaluation the evaluators constructed the chain of reasoning for Program B as shown in Fig. 3.

The model describes how an individual firm through participation in the program was expected to improve competitiveness and development environment (its ability to improve itself). The model describes the pro-

[3] The official Norwegian name of Program B: BedriftsUtvikling med Ny Teknologi (BUNT); in English: Business Development Using New Technology.

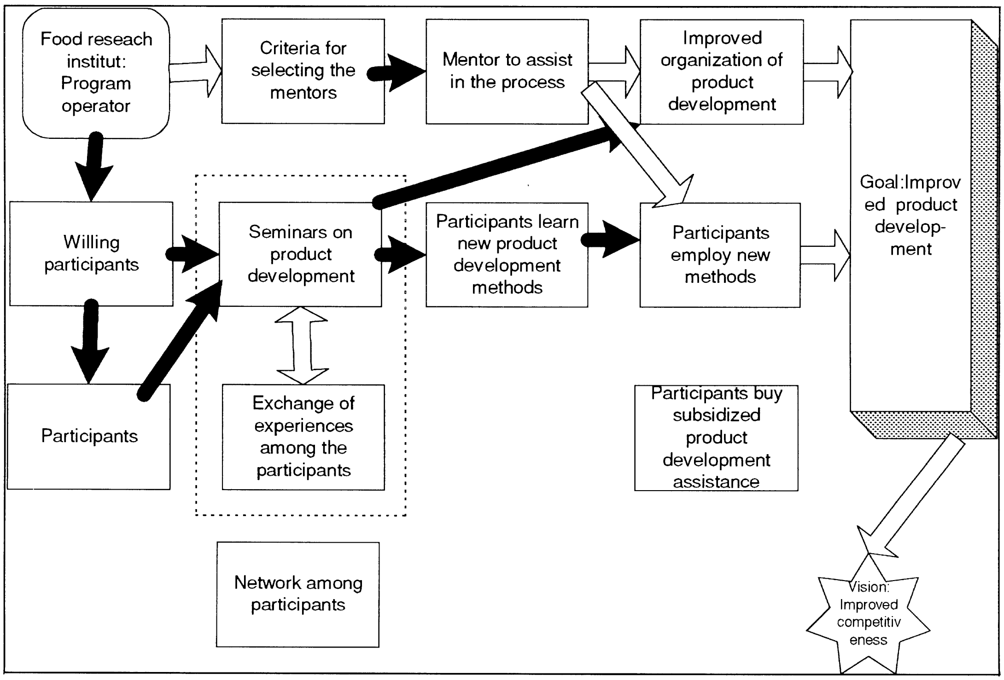

Fig. 2. Evaluation of pilot project of Program A: revised chain of reasoning (from Torvatn, 1996).

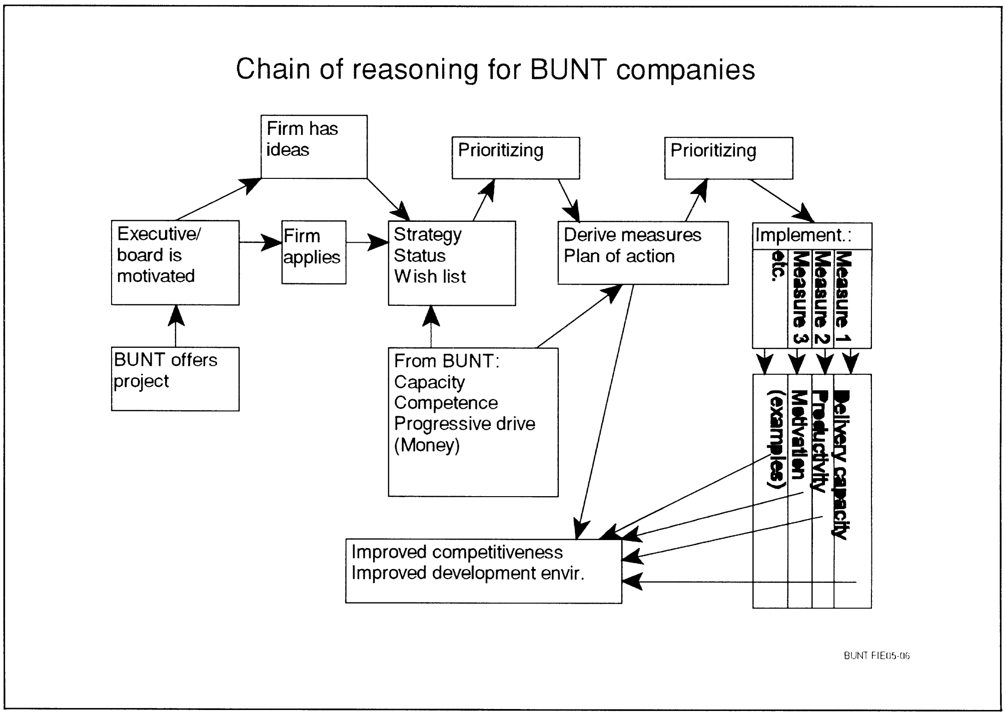

Fig. 3. Chain of reasoning for Program B (BUNT).

gram as seen from the point of view of one of the participants. The starting point of the process is the offer from the program of participation, and the ultimate goal is improved competitiveness. It will be seen that the model contains a heterogeneous set of activities, goals, external inputs, etc. The arrows may indicate both time and logical precedence.

This chain of reasoning provided an excellent tool for summative evaluation of the program. The evaluators simply followed the chain of reasoning and investigated how far the companies had progressed along the chain. Some of the progress was quite easy to track, like applications, development of strategy and an action plan. Of course some of the participants did not make it through the whole program, the overall drop out rate was 22%. However, the important question in the summative evaluation was whether or not the program had had any impact on those 78% completing. While it was difficult to measure competitiveness in the enterprises a short time after participation the evaluators analyzed how the companies had been able to implement their plans. Having copies of these plans, the measures and the prioritization it was possible to tailor a general survey of each participant, and ask each company whether or not each measure planned had been implemented.

Through this analysis the evaluators found that on average the companies had identified and prioritized between 5 and 6 measures.[4] A typical measure had an average prize of 38,000 USD, all of the companies completing claimed to have initiated implementation of at least one measure. Eighty nine percent had finished implementation of the most important measure, typically a company had implemented two to three measures a short time after completion of BUNT. After the first six months the likelihood of completion dropped quickly. Of all measures in all companies 39% were completed, 11% expected to be soon completed, 32% were initiated while 18% were not initiated. Overall, the evaluation concluded that the chains of reasoning of Program B had held so far, the program had been able to influence and to some extent trigger measures aiming at improving its competitiveness. The question still remains whether or not these measures have achieved the ultimate goal of improved competitiveness. This is a question that can only be answered over a long period of time. Some early results found a significant and strong correlation between first year improvement and degree of implementation of planned measures (Finne, 1994). Improvement was here measured as turnover per man-year, relative to the year of entry in the program. This research has however not been continued after the first year.

The problems in measuring the ultimate goal indicates that the chain of reasoning should have included some near-term indicators of competitiveness, inserted after the implementation. Examples of such indicators would be market entry or exit, market share in important markets or rate of innovation. Still, it would take some years before changes in these factors could be measured, and the time lag before measurement is possible and lack of continual post program evaluation remain the most important obstacles in summative evaluation of industrial modernization efforts.

[4] Average: 5.66 measures.

6. Program C[5]: chains of reasoning as a tool in additionality analysis

6.1. *Description of Program C*

Program C was one of the spin-offs from Program B. It was commissioned by the European Union (EU), and operated in 17 European countries, covering more than 1200 SMEs. The total EU funding amounted to approximately 6 million dollars. The core purpose of Program C was "promoting the absorption of new technologies by SMEs through the use of experienced consultants in the management of innovation". The basic idea in Program C was to train consultants in the use of a specific tool kit selected for the program. The consultants should then carry out Program C assignments in SMEs. In these assignments the consultant should conduct a strategic analysis of the company, identify technology which the company could profitably implement, and produce an action plan to assist the company in its implementation process. Through this process it was hoped that the companies would improve their competitiveness as well as learn what was called Integrated Management Techniques (IMT). Strategic planning and ability to use consultants were among these IMTs.

Since Norway had originally developed Program B a specific version of Program C was developed for Norway. The Norwegian program operator had an innovative idea for a Norwegian version of Program C: the program would recruit former Program B companies and test the efficiency and efficacy of reinforcing the strategy process in these companies through a five day reanalysis. This was accepted by the EU, and a total of 43 such assignments were carried out. Since this was an experimental scheme it was evaluated continuously, something not normally done with so small a program. This paper discusses the Norwegian program and its evaluation.

6.2. *Use of chains of reasoning in Program C*

The evaluators constructed the following chain of reasoning for Program C:

As can be seen from Fig. 4, the logic of Program C is similar to the logic of Program B (see Fig. 3). Having evaluated Program B earlier the evaluators also at outset judged this chain of reasoning as one with a fair possibility of success. It would also be easy to track the participants along the chain of reasoning, as had been done in Program B, and use level of implementation of measures as success criteria.

However, level of implementation of measures is only

[5] Official name of Program C: Managing the Integration of New Technology.

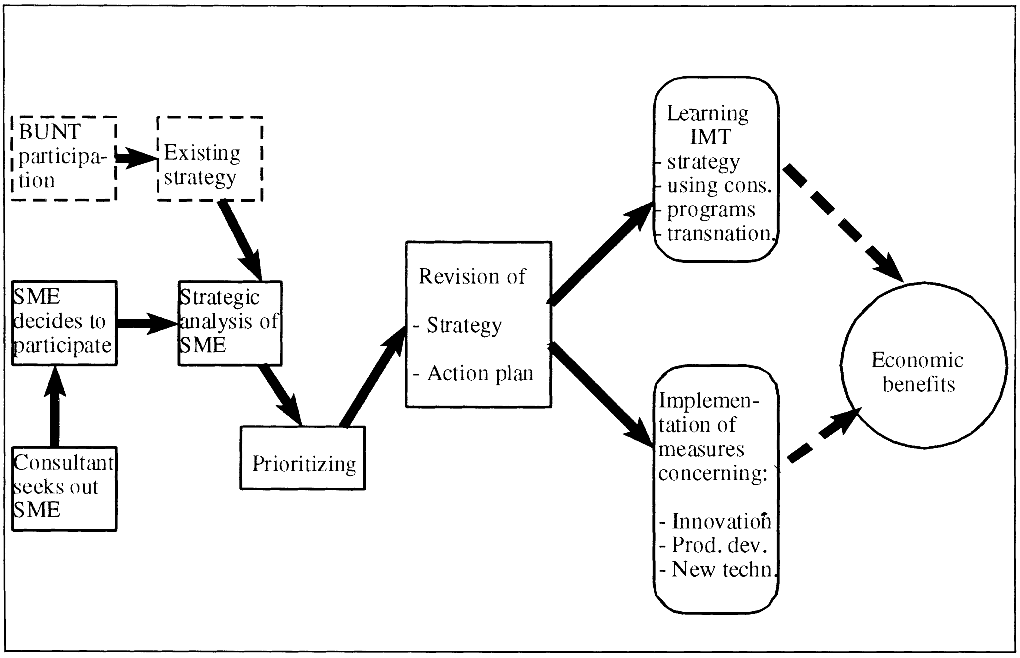

Fig. 4. Original chain of reasoning for Program C (MINT) (from Finne & Torvatn, 1996).

a valid criterion for success for a program if the program had some additionality. If no additionality could be found the program would simply be subsidizing existing plans. Additionality is one of the most important aspects of evaluation of industrial modernization (Finne & Torvatn, 1996; Georghiou, 1998), and is concerned with questions like: Did the intervention make a difference? What would have happened without the intervention?

In the context of industrial modernization additionality could be defined as follows: "Additionality is the difference between the benefits with which the program is associated and those that the company would have incurred with the most probable alternative use of its own resources" (Finne & Torvatn, 1996, p. 37). The question of additionality had been briefly discussed in the evaluation of Program B, but in the evaluation of Program C a more thorough analysis was carried out. The analysis focused on the following possible effects:

- *Timing* effects. This occurs when the company takes action and implements measures earlier than it otherwise would have done. This usually results in benefits accruing earlier, too, but premature implementation may also result in additional cost.
- *Scale* effects. This occurs when the company does more of a certain activity due to the public support. A particular case occurs when the public program makes things happen that would not have taken place at all, either because the company had wanted to do it but would not dare without the external support, or because the program brought something entirely new that the company accepted. This is called a *release* effect.
- *Directional* effects. With the use of program resources, including competence and program specific solutions, the company may end up doing things differently—and, hopefully, better—than without the program input.

Measuring additionality as a factual piece of information is always impossible because it has to deal with counterfactual data. The best one can do is to establish as credible indicators as to what possibly would have happened without the program.

Having the caveat in mind an additionality analysis was carried out. A release effect was found among 40% of the companies, who would not have done a similar project without Program C. Of the other 60% 20 did not know, while 40 said they would have done a similar project (no release effect). Other scale effects were not found. A timing effect was found in 57% of the companies, who would have postponed the project without support. Only 16% would have started the project at the same time. Thus Program C had had some release effect and some timing effect, whether the level of additionality was high enough was difficult to say, given the lack of standards to compare with. The evaluation concluded that the level was satisfying.

While scale and timing effects are relatively easy to identify directional effects may be harder to spot. From the chains of reasoning in Fig. 4 the evaluators could identify a key question concerning the additionality of Program C: how had the old plans from Program B been used in the strategic analysis and the following prioritization? To what extent could Program C provide new direction for the participants?

Three questions were used to provide an answer. The first was concerned with the newness of the measures, to what degree had the company participating already planned such measures? It was found that 47% had a loose idea about the project, 42% had started planning before Program C, while only 6% found a new measure and another 6% were able to 'restart' a measure earlier aborted. The second question was concerned with the role of the old plans from Program B: was there any follow up of measures from Program B? Eighty two percent of the companies answered yes to that question. The third question was related to the strategic planning process: would the companies have performed a strategic analysis without support from Program C? Fifty three percent of the companies answered yes to that, 34% no, and 13% did not know.

Based on these three questions, the evaluation concluded that Program C had had little directional effect. Rather the old plans from Program B had provided direction for the companies. A chain of reasoning was constructed summing up these and other results:

As can be seen from Fig. 5, Program C, to a great extent, functioned as planned. The consultants were able to recruit the desired number of companies in a rather competitive program environment, and all recruited companies completed. The measures were implemented at approximately the same rate as in Program B, and participants knew how to use some new management techniques like strategic analysis and use of consultants. They had probably not learned this through Program C, however, since they all had such knowledge already.[6] Further, the prioritization had not worked as planned. The participants had used items from their old plans in most of the cases. Hence the filled arrow from the old BUNT (Program B) to the revision, noting a strong influence, while the new analysis only had a small and limited influence.

7. Discussion

Above three cases of evaluation using chains of reasoning have been presented. The cases have described various forms of use of these chains, Table 1 below sums up these forms, as well as other forms of use of chains of reasoning not discussed in this paper:

As described in the case of Program A, chains of reasoning are excellent for describing program logic and doing an evaluation of the concepts underlying the program. This is also a good starting point for identifying key issues for investigation, that is, focusing the evaluation.

Impact analysis, or rather to reduce the problem of time lag when conducting impact assessment, was the reason chains of reasoning were developed. The underlying logic is simply that as long as the program keeps on track possibility for impact is increased, as in the case of Program B. The evaluator must, however, remember to

[6] Data on this is not presented here; see Finne and Torvatn (1996) for details.

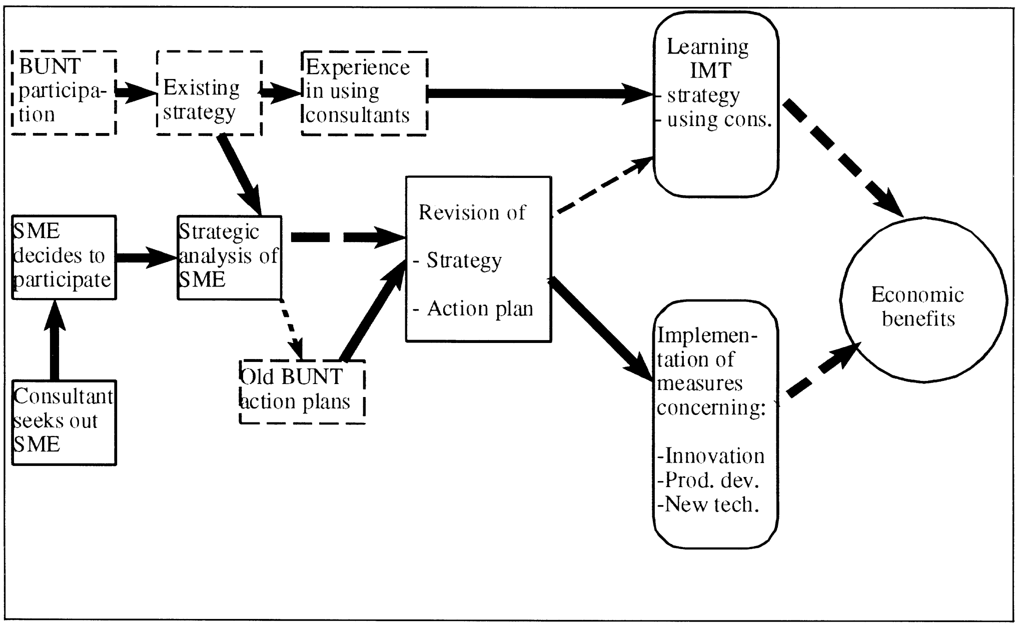

Fig. 5. Revised chain of reasoning for Program C, derived from the evaluation (from Finne & Torvatn, 1996).

Table 1
Forms of use of chains of reasoning

Form of use	Case
Description of program logic-concept evaluation	Program A
Identifying key issues	Program A, Program C
Impact assessment	Program B
Communicating findings	Program A, Program C
Monitoring	Not discussed here
Mapping different program theories among different stakeholders	Not discussed here
Focusing the evaluation	Program A

check for additionality. As the case of Program C demonstrates, a program can keep on track but still have limited contribution.

An important aspect of chains of reasoning in impact assessment is that they enable the evaluator to start doing impact assessments from day one. The monitoring aided by chains of reasoning tells the evaluator whether or not the program is on track. If steps in the chain are not carried out, or fails, the possibility for impact is reduced (frequently to zero).

Chains of reasoning are also excellent tools for communicating evaluation findings. A drawing (with a brief description) is easy to understand and can convey the main findings in a non-technical language to a large audience.

Finally chains of reasoning could be used in monitoring progress and identifying different stakeholders program theories on the program in question. These forms of use are not discussed here; see Torvatn (1995).

While chains of reasoning are flexible and useful in many different evaluation settings, they have definite limitations and should not always be used. For one thing, when evaluating agencies and institutions with one goal and several measures for reaching that goal the evaluator might simply be bogged down in a multitude of chains of reasoning. Chains of reasoning works best in a program evaluation, where the same sets of actions are repeated several times.

Chains of reasoning are also sometimes too rationalistic. It is assumed that it is possible to construct rational causal sequences of means and ends. This construction process is in itself problematic. The evaluator might ascribe certain goals they do not really have to persons and organizations, and also the evaluator will rely on theories espoused of stakeholders rather then theories in use. This is so because espoused theories are better documented and are intended for presentation, and thus easier to model.

Chains of reasoning also simplifies complex processes. This is one of the strengths of the tool, but it is easy to simplify too much. In order to compensate for this it is of course possible to construct overly complex drawings, thus destroying the simplicity that is one of the tools foremost virtues. Furthermore chains of reasoning have difficulties in incorporating the effects of context on processes. And finally chains of reasoning takes little advice of power relations, of symbols, rituals, culture and all sorts of 'non-rationale' activities. As any tool chains of reasoning has its limitations.

However, as this paper has attempted to demonstrate, chains of reasoning are easy to use, flexible, and useful for several purposes. Even with the limitations described above, chains of reasoning should be a valuable tool in any industrial modernization evaluators tool-kit.

References

Arbo, P. (1993). *Teknologi og kompetanseorientert støtte. En oversikt over utvalgte tiltak og evalueringer*. (Technology and competence enhancing efforts. An overview of selected efforts and evaluations.) (SF037-4138). Tromsø: Norut Samfunnsforskning.

Chen, H. (1990). *Theory driven evaluations: A comprehensive perspective*. Newbury Park, California: Sage.

Chen, H., & Rossi, P. (1983). Evaluating with sense—the theory driven approach. *Evaluation Review*, *7*, 283–302.

Chen, H., & Rossi, P. (1987). The theory driven approach to validity. *Evaluation and Program Planning*, *10*, 95–103.

Chen, H., & Rossi, P. (Eds.). (1992). *Using theory to improve program and policy evaluations*. (Vol. 290). Westport, Connecticut: Greenwood Press.

Finne, H. (1994). How does strategic planning influence technology strategies in small firms? In A. Francis, S.-Å. Hörte, & J. L. Pedersen (Eds.), *Designs, networks and strategies* (Vol. 1, pp. 172–196). Luxembourg: European Commission, Directorate-Generale XIII.

Finne, H., Levin, M., & Nilssen, T. (1993). *Strategisk bedriftsutvikling på norsk. Evaluering av BUNT-programmet*. (Strategic business development in Norwegian. evaluation of the BUNT-program.) (STF82 A 93001). Trondheim: SINTEF IFIM & ORAL-NTH.

Finne, H., Levin, M., & Nilssen, T. (1995). Trailing research: a model for useful program evaluation. *Evaluation*, *1*(1), 11–31.

Finne, H., & Torvatn, H. (1996). *Evaluation of MINT in Norway* (STF38 A96514). Trondheim: SINTEF IFIM.

Georghiou, L. (1995). Assessing the framework programmes—a meta evaluation. *Evaluation*, *1*(2), 171–188.

Georghiou, L. (1998). Issues in the evaluation of innovation and technology policy. *Evaluation*, *4*(1), 37–51.

Rosenfeld, S. A. (1996). Does cooperation enhance competitiveness? Assessing the impacts of inter-firm collaboration. *Research Policy*, *25*(2), 247–263.

Rothwell, R., & Dodgson, M. (1992). European technology policy evolution: convergence towards SMEs and regional technology transfer. *Technovation*, *12*(4), 223–237.

Rush, B., & Ogborne, A. (1991). Program Logic Models: expanding their role and structure for program planning and evaluation. *The Canadian Journal of Program Evaluation*, *6*(2), 95–106.

Shapira, P., Youtie, J., & Roesner, J. D. (1996). Current practices in the evaluation of U.S. industrial modernization programs. *Research Policy*, *25*(2), 185–214.

Torvatn, H. (1995). *Chains of reasoning*: *An evaluational tool*. (IFIM Paper 5-95). Trondheim: SINTEF IFIM.

Torvatn, H. (1996). *Evaluering av pilotprogrammet i Matforsks PU-Nettverk* (STF 82 F95013). (Evaluation of the pilotproject in the Food Research Institutes Product Development Networks) Trondheim: SINTEF IFIM.

Torvatn, H. (1997). *En analyse av utvalgte programmers betydning for å stimulere tilꝃkt produktutvikling i bedrifter*. (An analysis of a selection of programs ability to influence product development in businesses.) (STF38 A96523). Trondheim: SINTEF IFIM.

Torvatn, H., & Elvemo, J. (1997). *Evaluering av MATFORSKs PU-program-en del av konkurransestrategier for norsk mat* (STF38 A97506). (Evaluation of the Food Research Institutes Product Development Networks- part of the 'Strategies for competitiveness of Norwegian food and food producing industry effort') Trondheim: SINTEF IFIM.

Torvatn, H., & Munkeby, I. (1994). *En kartlegging av norske tiltak for n*œringslivet. (STF82 A94010). (A survey of Norwegian industrial modernization efforts.) Trondheim: SINTEF IFIM.

Weiss, C. (1996). *Theory-Based Evaluation*: *Past*, *Present and Future*. Paper presented at the American Evaluation Association Annual Meeting, Atlanta, Georgia, 8 November.