Pilot evaluation in TENCompetence

Pilot evaluation in TENCompetence: a theory-driven model1

Judith Schoonenboom*, Henk Sligte, Ayman Moghnieh, Marcus Specht, Christian Glahn, Krassen Stefanov

*SCO-Kohnstamm Institute, University of Amsterdam, Nieuwe Prinsengracht 130, Amsterdam, NL-1018 VZ, The Netherlands

E-mail: judith.schoonenboom@uva.nl

Abstract: This paper describes a theory-driven evaluation model that is used in evaluating four pilots in which an infrastructure for lifelong competence development, which is currently being developed, is validated. The model makes visible the separate implementation steps that connect the envisaged infrastructure at the very beginning to the actual learning outcomes at the end. The model shows how each implementation step can lead to differences, that may influence the ultimate outcomes. We first show the characteristics of the model, and how they are derived from existing evaluation literature. Secondly, we show how the model is used in the several stages of evaluation, including its use to test the program theory.

Keywords: lifelong learning, competence development, evaluation model, multisite evaluation

Introduction

This paper is about the design of the pilot evaluation in the TENCompetence project. The TENCompetence project is a four-year project in the European Commission's 6th Framework Programme, priority IST/Technology Enhanced Learning. The aim of the project is to design a technical and organizational infrastructure for lifelong competence development. The project develops new innovative pedagogical approaches, assessment models and organisational models, and it creates a technical and organizational infrastructure which integrates existing isolated models and tools for competence development into a common framework.

The TENCompetence infrastructure will be validated in four different pilots, representing the variety of contexts in which lifelong competence development takes place. One pilot in Spain aims at lifelong learners in the film industry who have to learn to master digital cinema techniques. One pilot in Bulgaria aims at teachers who have to update their skills in using ICT in teaching. A third pilot in several African countries aims

[1] The work on this paper has been sponsored by the TENCompetence Integrated Project that is funded by the European Commission's 6th Framework Programme, priority IST/Technology Enhanced Learning. Contract 027087 (http://www.tencompetence.org/)

at students in water management, and a fourth pilot in Belgium aims at establishing an infrastructure for Antwerp as a lifelong learning city. The evaluation of these pilots is a complex undertaking, caused by expected large differences among the pilot sites. Learning objectives will be different, the intended audience and thus the learners involved will be different, and as the system is under development, the system implementation and the way in which the system is used, will be different.

To get and hold a firm grip on such a complex evaluation, a solid evaluation design is an essential requirement. In this paper we describe a model that was developed for underpinning the evaluation design, and we explain how the model is and will be used in designing the evaluation processes.

The model

<Figure 1 goes here>

Figure 1 The pilot evaluation model, showing the steps between envisaged infrastructure and ultimate outcomes

Figure 1 presents our model of the pilot evaluation. In short, the model presents the TENCompetence infrastructure, as envisaged by its developers at the beginning of the project (entirely to the left), and the ultimate outcomes of learners who have used the infrastructure in the pilots (entirely to the right). In between, a hierarchy of steps is shown that represent the progressing steps in implementing the infrastructure. In research terms, these steps represent a hierarchy of variables that possibly influence the outcomes (Costner, 1989). The implementation actors are in succession: the developers of the infrastructures, the providers of competence development programs and the learners. The components of the model will be discussed in more detail below.

The model is intended to guide the pilot evaluation in two ways. Firstly, it enables the description of the pilot results at different levels, such as ‘how was the infrastructure implemented’, ‘how was it used’, ‘what were the results of use’. Secondly, by comparing the envisaged use and learning outcomes to the actual use and outcomes, the model can be used in testing the program theory, i.e. the hypothesized relations between infrastructure and outcomes.

Model components

This section describes six characteristics of the model in more detail and links these to the literature. Firstly, the model is a from of implementation evaluation (Patton, 1997), in which the chain of steps between vision and implementation is described (King, Morris, & Fitz-Gibbon, 1987; Williams & Elmore, 1976). The following chain is set up: in step one, the envisaged infrastructure is materialized in a technical implementation of the infrastructure, consisting of various pieces of software with different functionalities. In step two, models for lifelong competence development are provided that make clear how the infrastructure can be applied for lifelong competence development. These two steps are performed by the developers of the TENCompetence infrastructure. In step 3, providers of competence development programs in the pilots make a selection out of the available technology and models that fit the needs of their intended audience. In step 4,

learners actually use the materials, models and tools that are made available to them by the pilot providers. Finally, in step 5, out of their use, learners get their experiences and produce specific learning outcomes.

A second characteristic of the model is that it focuses on discrepancies between the ideals and the actual outcomes that result from the implementation process (Provus, 1971). For example, according to the TENCompetence pedagogy for lifelong competence development, learners should be enabled to select their own competence development program. This implies that the route along which learners reach specific learning objectives may well be different for different learners. In the technical infrastructure of step 1, this possibility of choosing one's own competence development program has been implemented. However, in step 3, it is conceivable that a provider offers 'one size fits all' competence development programs, which do not allow learners any freedom. In that case there is a discrepancy between the available infrastructure in step 2 and the infrastructure offered by providers in step 3. Discrepancies may arise at other places, for example between the infrastructure offered by providers in step 3 and the use that learners in step 4 make of this infrastructure. If the providers enable sharing of resources between learners, but learners do not use that functionality, then there is a discrepancy between these two stages.

Thirdly, the model focuses on differences between pilot sites. In fact, the discrepancies mentioned above are a special type of differences. These are differences that result from concepts in a specific step not being implemented in the next steps. But there are also more general differences between pilot sites. For example, the model includes differences between learners. Learners may differ in their background and learning needs. And when compared on the content levels, the resources used in separate pilots can be very different. More generally, all variables that are expected to influence the outcomes, are included in the model, rather than 'controlled' by randomization (Chen & Rossi, 1987; Costner, 1989).

Fourthly, the model focuses on outcomes, rather than services delivered. The learning outcomes are explicitly included in the evaluation (Patton, 1997).

Fifthly, the model includes the needs of the learners (Scriven, 1993). In the model, learner needs are included as a property of the learner. In the evaluation, they have two functions: firstly, differences in learner needs between learners and between the pilots are measured and are considered to be part of the relevant differences in pilot implementation, which may account for differences in outcomes. Secondly, learner needs are used in measuring the outcomes. Part of the outcomes includes the extent to which the learning experience using the infrastructure meets the learner's needs.

Finally, the model focuses on the program theory and on testing this theory. The envisaged use of the infrastructure and the envisaged outcomes are part of the model (entirely to the left). Also included are the envisaged mechanisms that relate the use of the infrastructure to the outcomes. This is called the program theory, and the corresponding evaluation is called theory-driven evaluation (Chen & Rossi, 1987; Patton, 1997). These are the ideas of its inventors on how they think the TENCompetence infrastructure works. Part of this theory are arguments such as 'if learners can determine their own learning path, the units of learning chosen will better match their need and if chosen units of learning better match their needs, learners will learn more' or 'if prior learning is assessed, then units of learning can be chosen that better match the learners proficiency level; if chosen units of learning better match the learners' proficiency level, learners will learn more'.

Author

The outcomes, mentioned under point four, play a central role in testing the theory. An important part of the evaluation is testing whether the program theory holds in the pilots. If, for example, at a particular pilot site, learners cannot choose their own learning path, but yet do not learn less than learners who can choose their own learning path, then obviously, there is a problem with the program theory of the last paragraph.

Model use

The model has played an important role in shaping the pilots and the pilot evaluation, and it will continue to do so. To start with, the model has helped in setting-up the pilots. For many of the model's elements, lists are available that list relevant possibilities. These lists have been used not only in clarifying the intended set-up of the pilot, but occasionally, this has also led to a change in the pilot set-up. For example, one of the step 2 TENCompetence models in development is the pedagogical model for lifelong competence development. This model contains a list of pedagogical requirements that an infrastructure for lifelong competence development should meet. One such requirement is that in lifelong competence development teacher and assessor are separate roles. As competence assessment aims at assessing learning in all learning contexts that the learner encounters, including the competence development program followed by the learner, the assessor will often not be the teacher, as there often will be no teacher. The idea is also to separate the two roles to guarantee an independent judgement. At the beginning of the pilot set-up, this feature of the pedagogical model made it clear to the pilot coordinator of the first pilot, that in their pilot, the role of teacher and assessor was not separated: teacher and assessor were the same person. Furthermore, after reflection, the pilot coordinator decided to separate both roles, to guarantee a more independent judgement.

The model helps to clarify, and therefore to handle, differences between different stages of implementation and differences between pilot sites. Imagine that a model would only consist of the envisaged infrastructure on the one hand and the learning outcomes on the other hand. Compared to Figure 1, such an evaluation model is a very impoverished model. Yet, all to often program evaluations more or less use such an impoverished model, distinguishing only between 'treatment' and 'outcomes'. However, such an impoverished model mixes up differences between several steps of implementation, thereby increasing the error considerably. Furthermore, in a multisite situation, an elaborated model as in Figure 1 is needed to pinpoint the differences and thereby make them manageable. Simply stating that the implementation at different sites will be different, is not enough, and will make comparative analysis very hard.

Currently, the model plays an important role in clarifying the program theory. Above, we presented some of the propositions that we suppose are part of the program theory. Up till now, the program theory and its propositions have mainly been derived from the formulation of the TENCompetence objectives and solutions in the TENCompetence domain model and in the project's description of work. Currrently, these propositions are being checked with the initiators of the project. One purpose of the evaluation is to validate the program theory, and therefore we must be sure that we are testing the right theory.

Another major role of the model is in focusing the evaluation. As the implementation of the elements of the model can be done in so many ways, elaboration of the model leads

to a large list of opportunities for evaluation. Having this list is a first step in focusing, which will consist of selecting the most important possibilities from the list.

Finally, the model helps in looking for comparison with a 'control group' in a way that makes sense. Obviously, it doesn't make sense to compare a group of people who follow a competence development program using the TENCompetence infrastructure to a comparable group of people who do not follow such a program. Of course, the program will have an effect: it would be very strange if people worked on developing a competence for a while, and this does not have any effect (Patton, 1997). The model provides clues to useful comparisons. It makes clear that it is the intended ways in which the infrastructure is used, based upon the underlying models of step 2, that are supposed to make a difference. One way of testing this, is to compare a group of learners who use the infrastructure as intended, and one group who use the infrastructure, in which the TENCompetence models have not been implemented. For example, one may compare learner who are enabled to choose their own learning path, versus learners who cannot do this.

Conclusions

In this paper, we have presented a theory-driven model for multi-site pilot evaluation. We have demonstrated features of the model and we have shown how the model is used in designing the pilot evaluation. We have shown that the features of the model, although derived from the theory of social program evaluation, can be very useful in evaluating a technical and organizational infrastructure for lifelong competence development.

References

Chen, H.-T., & Rossi, P. H. (1987). The theory-driven approach to validity. *Evaluation and Program Planning, 10*(1), 95-103.

Costner, H. L. (1989). The validity of conclusions in evaluation research: A further development of Chen and Rossi's theory-driven approach. *Evaluation and Program Planning, 12*(4), 345-353.

King, J. A., Morris, L. L., & Fitz-Gibbon, C. T. (1987). *How to assess program implementation* (2nd ed ed.). Newbury Park, Calif.: Sage Publications.

Patton, M. Q. (1997). *Utilization-focused evaluation : the new century text* (3rd ed ed.). Thousand Oaks, CA: SAGE.

Provus, M. (1971). *Discrepancy evaluation : for educational program improvement and assessment*. Berkeley, CA: McCutchan.

Scriven, M. (1993). *Hard-won lessons in program evaluation*. San Francisco: Jossey-Bass.

Williams, W., & Elmore, R. F. (Eds.). (1976). *Social program implementation*. New York etc. ; London: Academic Press.

Author

Figures