TEACH AS YOU PREACH -

LESSONS FOR PROCESS-ORIENTED EVALUATION OF TRAININGS

TEADE PUNTER[1], DANILO ASSMANN[2], DIEUWKE DE HAAN[3]

Abstract

This paper presents a process-oriented method to evaluate training approaches for continuous job-oriented education. The method is designed in order to conduct economically feasible evaluation of training approaches, meanwhile not focusing on the increase of participants' competences, but instead on the training processes.

To apply the ISO 15504 philosophy to this area, a process reference model for job-oriented education has been developed, the assessment model has been adapted, and was applied for Level 1 capability ratings.

Our evaluation method was successfully applied within a large German Telecom company, which helped us to determine the strengths and weakness of the training approach as well as to come up with a set of improvement suggestions. Based upon these experiences we reflect on the possibilities to apply this method when training SPiCE assessors.

Keywords: Process-oriented evaluation, continuous on-the-job education, constructivist learning approach, base practices, ISO 15504.

1   Introduction

In today's educational market for corporate trainings there is an increasing amount of training approaches that give room for self-directed, active and experience based learning (Simons et al, 2000). These process-oriented approaches find their roots in the constructivist view of learning that holds that learners are learning while conducting authentic problem solving tasks that are situated in real-life environments. Learners have to set their own goals, control their own learning process and follow their

---

[1] Embedded Systems Institute, the Netherlands
[2] Valeo Interior Controls, Germany
[3] Eindhoven University of Technology, the Netherlands

own learning strategies. Meanwhile they are stimulated to reflect upon their own learning process (Kirschner, 2000; Gerstenmaier and Mandl, 2002).The quality of learning depends on the quality of the processes that should take place. Examples of training approaches that have applied the constructivist view on learning are *Virtual Enterprise* that was developed in the mid 90-ties in the Netherlands (Van Petegem, 2000) and *Working-process-oriented education*[4] that was developed in the beginning of this century in Germany.

Because constructivist learning approaches assume that the quality of learning depends on the quality of the processes that should take place, the focus of evaluating the associated training approaches should be on processes too. A process-oriented evaluation method – an approach that focuses on analyzing the learning activities as well as the interactions between the roles in the learning process – is therefore appropriate. Our approach for process-oriented evaluation of training approaches is inspired by ISO 15504/SPiCE (ISO, 1998). This method is – like the Capability Maturity Model (Paulk, 1995) – meant for software processes or organizations, and it has proven to conduct effective evaluations of the quality of the software development processes (El Eman et al, 1998).

This paper discusses the theory and practice of a process-oriented evaluation method for training approaches. The method is described in section 2. Section 3 presents the results and experiences in applying the method for evaluating the APO learning approach at a German Telecom company.

2   The Method for Process Evaluation

This section describes the evaluation method that is to be used for the evaluation of the process quality of process-oriented training approaches.

2.1 Evaluation goal

The goal of the evaluation is to enable the improvement of the training approach when insufficient quality is detected. The evaluation should identify a set of strengths and weaknesses that can be used to start an improvement cycle of the training concept.

2.2 Evaluation object

The object of evaluation is the process-oriented training approach that will be inspired by a constructivist view of learning. A training approach is defined as a set of guidelines that 1) describe how to conduct training activities, 2) identifies the people (roles) for the training, 3) provide documentation rules for the participants to write down and reflect upon their learning experiences and 4) define how the quality of the training approach has to be assured.

---

[4] APO: "arbeitsprozessorientierte Weiterbildung"; see: www.apo-it.de

The training approach is implemented into specific trainings by applying the guidelines within a particular training – e.g., for network administrator – at one or more company department(s). The consequence for the evaluation is that the information about the evaluation object is collected from the implemented trainings.

The data about these trainings are reported by documenting the implementation and collecting the experiences of particular roles, like participants, see section 2.4, that are involved in a training. It is assumed that each of the organizers implements the training according to the training approach.

The challenge of evaluating a training approach is to generalize the information that is collected by interviewing the people that play the different roles during a training upon an aggregation level that will provide us with reliable information about the training approach itself. For this purpose, appropriate interview partners have to be selected which are involved in representative trainings by different organizers, see Figure 1.

The selection is done by looking at the degree to which the implemented trainings and participants represent the training approach. This means that more training implementations have to be consulted in case of a higher variety of implemented trainings – e.g., not only a training for network administrator, but also for IT Systems analyst, Software developer and/or in different organizational settings. The ambition level of the training approach is another factor that influences the selection of trainings and participants. When a training approach claims to result in major improvements, compared to participants' original situation, than this means for the evaluation that more implementations have to be consulted to verify the improvement.
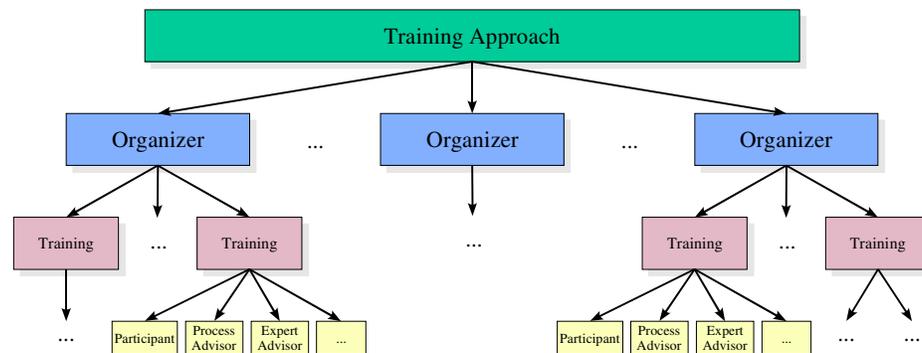
.

Figure 1: Evaluation Object Architecture: implementation of the training approach into training measures that are applied by participants.

2.3 Criteria: base practices

The evaluation of a training approach requires process-oriented criteria instead of focusing on product-oriented criteria like the competences or the skills of the people. In fact, a reference educational process is needed that can be compared to the training approach under evaluation. The difficulty in finding references for training concepts is that each implementation of the training approach leads to unique learning ar-

rangements for the participant, i.e. a unique set of processes. Therefore, the evaluation of the training approach should be to cope with the diversity of process implementations (the trainings).

A effective way of describing reference processes is the approach of base practices that is applied for assessing software processes (ISO 15504, 1998). A base practice describes what needs to be done by the (sub-) process in order to achieve one or more goals. Base practices address the infrastructure as well as the activities of an instantiated training process, but they are formulated according to a more general pattern. For example, they address only the 'what' has to be done, never the 'how' it has to be done.

The base practices were defined by a 3 step approach following the ISO 15504 logic, starting with four process areas, see also Table 1:

- Planning – the planning activities to start the training approach,
- Preparation – the preparation activities to start the execution of the training approach,
- Execution – the actual performance of the training approach,
- Wrap-up – the certification of participants and the ending of the training approach.

The base practices are structured according to three refinement steps. The 1st refinement denotes the phase of the training process (planning, preparation, execution and wrap-up), which is in SPiCE terminology the process area. The 2nd refinement describes the processes (SPiCE terminology) per phase, for example the five processes of the Execution phase; see Table 1.

Planning	Preparation	Execution	Wrap-up
Documentation of knowledge, skills and roles (3)	Define training plan (1)	Competence and availability of learning and information support (8)	Qualification certificate (3)
Characterization of training environment (5)	Provide communication interfaces (6)	Process oriented execution (9)	Internal evaluation (3)
Determine initial situation (starting point) (3)	Provide adequate Environment (3)	Quality of training (4)	External evaluation (1)
Define goals (5)	Assess and qualify process advisor and expert advisor (6)	Quality assurance measures (4)	Services (2)
Select and decide about training (2)	Provide advice to participants about training (1)	Provision of adequate learning material and infrastructure (8)	
	Documentation of training/training approach (4)		

Table 1: Overview of processes per process area (numbers in parenthesis indicates the number of base practices)

The 3rd refinement is about the activities or infrastructure issues that apply to the processes, which are the base practices in SPiCE terminology. The base practices

describe the requirements what is supposed to be done. This can include infrastructure issues and other constraints or prerequisites for the processes.

The evaluation is based on the 3rd level of refinement, which enables the evaluation because for each of the process it can be checked if it is achieved by the learning approach. Such conformances check aims at determining how far a process is covered by the training approach. For example, to measure the coverage of the process "Quality of training", which is a element of the Execution phase" it has to be checked if the following four base practices are realized:

- Learning goals are taken into account,
- Interest areas of participants are regarded,
- The training(methods) are defined in a way that they are balanced to the number and ambitions of the learning goals,
- The number of participants should be appropriate for the learning approach.

Each of the base practices within the process addresses one or more roles. For example, the statement that "Learning goals are taken into account" has to be verified by three roles: Participants, Process advisor and Expert advisor.

According to this three level refinement structure, a set of 20 processes was defined. These processes are organized according to 4 process phases (also process areas) and can be refined in a total of 77 base practices. The number of processes and base practices was determined by considering that a large amount of processes is needed to get a detailed view, meanwhile noticing that their number should be small enough to be still high-level and keep the overview.

2.4 Information sources: roles

In order to get the information about the particular trainings people should be interviewed. Each of those interviewees will play a specific role in the training. The following roles are distinguished:

- Participant – the individual that does the training and gains competence (this role is always a real person)
- Process advisor/mentor – this role leads the participant through the education process and also focuses on personal development (this role might be performed through a person or a system). The major importance of this role is that it provides the interface to the training approach for all other roles. The process advisor gives all learning support (e.g., methods and techniques) to the participant and provides also the organizational learning environment.
- Expert advisor – this role is the source for knowledge (can be a person or a system)
- Superior – this is the superior of the participant; this role defines (additional) goals for the education and demands/assesses a specific quality of the training

3 A Case Study on Evaluating a Training Approach

This section reports about applying the process-oriented evaluation method for training approaches that was presented in previous section. The evaluation was conducted

during a pilot of implementing the APO training approach –- within a German tele-communication company. The goal of the evaluation was to evaluate the quality of the APO continuous education. Risks of the training approach had to be detected to improve the approach and avoid those risks in an improved training approach.

3.1 Interviews

The pilot of the APO training approach was implemented with 14 participants working for the same company in different departments spread over Germany. Five process advisors were guiding those 14 participants. Four of the five process advisors and ´their` associated participants (also four) were selected for the evaluation, by arguing that with those people the broadest set of implementations of the training approach in the company was available.

These participants and process advisors were each interviewed by two evaluators to collect information about their trainings because it was thought that the interviewed people would show how the trainings were lived, and how good the lived process were covering the base practices. Each participant and process advisor was interviewed separately two times, namely the first time after the phases Planning and Preparation were finished and the second time after completing the APO approach. Expert advisors and supervisors were not selected for interviews, because of the restricted resources. In total sixteen (4 + 4) x 2 face-to-face interviews were conducted.

During and closely after finishing the interviews the extend to which the training process covers each of the base practices was rated independently by the two evaluators according to three categories: '-' means it is not considered in the actual process; 'o' means it is considered, but fulfils not all requirements; '+' means it is completely fulfilled. Furthermore, the statements of the interviewees were recorded, which resulted in additional feedback that was structured according to the base practices.

The ranking started with analyses of the interview reports per couple. The answers given by each participant on the questions about the base practices were compared with the answers given by the process advisor. Based upon these reports, each evaluator ranked independently for each base practice the coverage of the base practices according to the following ranking categories:

- FS: Fully satisfied – 100-85% of the base practices comply to the training approach,
- S: Satisfied – 85-50% of the base practices comply to the training approach,
- PS: Partially satisfied – 50-15% of the base practices comply to the training approach,
- NS: Not satisfied – 15-0% of the base practices comply to the training approach.

After the evaluators set their individual scores, they compared these scores to each other, which result in joint rankings per couple participant-process advisor. Discrepancies in scorings – e.g., between Satisfied and partially satisfied – were discussed by looking at the interview-protocols and then a decision on the score was made.

A scoring problem became visible during the first interview round, because the processes were defined too abstract and it was hard to score the processes. A compli-

cating factor was that the team of interviewers changed during the project, which made understanding of scoring the processes/base practices more difficult. However, during the second interview round the evaluators reached an agreement level of 70% for the base practices.

The scores per couple were aggregated by the evaluators into an overall score of the four couples per baseline. The result is an evaluation profile that expresses the extent to which the training approach – implemented by the particular training in this company – confirms to the base practices. Figure 2 presents an example of an evaluation profile, which is for reasons of confidence not identical to the APO training approach profile. The training profile shows the aggregated ranking of the processes based on the base practice ranking. Thus, the approach follows completely the approach of an Level 1 SPiCE assessment. The provided result is comparable with the level 1 process attribute.

Because, the scores of some of the base practices differed amongst the interviewed couples their conformance could not be set at once. The overall scores could be set for 50% of the base practices at once in this case study. In Figure 2 the effected processes of these base practices are depicted by the rows that do cover only one column of the evaluation profile (the blue areas). For the other 50% of the base practices there was no consensus at once (yellow bars). It was not possible to determine an overall score directly, because the scores of the interviewed couples were ranked in two adjacent ranking categories. Note: the bandwidth of the scores of the interview couples was never broader than two adjacent ranking categories. Post questionnaires were used to find a more unified score for these base practices, see next section.

Planning	FS	S	PS	NS
Documentation of knowledge, skills and roles	■ (blue)			
Characterization of training environment			■ (blue)	
Determine initial situation		■ (blue)		
Select and decide about training			■ (blue)	
Preparation	**FS**	**S**	**PS**	**NS**
Plan training		■ (blue)		
Provide communication interfaces		X (yellow)		
Provide adequate environment			■ (blue)	
Assess and quality process advisor and expert advisor	X (yellow)			
Provide advice to participants about training	(yellow)	X		
Documentation of training or training concept				
Execution	**FS**	**S**	**PS**	**NS**
Competence and availability of learning and information support		■ (blue)		
Process oriented execution		■ (blue)		
Quality of training		■ (blue)		
Quality assurance measures			■ (blue)	
Provision of adequate learning material and infrastructure		(yellow)	X (yellow)	
Wrap-Up	**FS**	**S**	**PS**	**NS**
Qualification certificate			■ (blue)	
Internal evaluation				■ (blue)
External evaluation		(yellow)	(yellow)	
Services	■ (blue)			

■ (blue)   Direct consensus between the four interviewed couples about the conformance of the training to the base practice

□ (yellow)   No direct consensus between the four interviewed couples about the conformance of the training to the base practice

X   The degree of conformance of the training to the base practice determined after further analysis

Figure 2: Example of an evaluation profile

3.2   Pre- and post questionnaires

Pre- and post questionnaires were organized to address more people than only the interviewed persons. Aim was to get extra data about the training approach, which was independent from the interview data.

The pre- and post questionnaires were derived from the base practices. They also covered the context, expectations, and elements of the organizational culture. Post-questionnaires were partly a crosscheck of the pre-questionnaires and also included questions that should solve issues that were still open after analyzing the interview scores. They were *not* used to test knowledge.

Questionnaires were sent to all people covering the roles in the pilot. The pre-/post-questionnaires consisted of 14/12 questionnaires for the participants, 5/5 to the process advisors, 11/11 to the superiors, and 12/12 to the expert advisors. The first number denoted the number of pre-questionnaires and the second number the number of post-questionnaires sent. The numbers show that two participants quit during the training (for personal reasons). The numbers above are the numbers of the questionnaires that were sent. We received 10/8 from the participants, 5/4 from the process advisors, 6/4 from the superiors, and 4/5 from the expert advisors.

The pre- and post questionnaires outcomes were used as a means to check the correctness of the overall scores found during the interviews that expresses the conformance of the APO training approach to the base practices. Especially the post questionnaires were beneficial when dealing with the fact that for 50% of the base practices no consensus on overall scores was found after our first calculation, see 3.1. The post-questionnaire enabled us to rephrase the questions about the base practices and to ask the different roles in the pilot again. These new and often less ambiguous formulated questions (than the original ones) resulted into less diversity in answering. In the end about 20% of the discrepancies remained, so a considerable reduction (from 50% to 20%) has been achieved with applying the questionnaires. The overall scores resulting from this step are presented as crosses (X) in the evaluation profile, see Figure 2.

3.3   Reporting

The reporting on process-oriented evaluations should present the strengths and weaknesses of the training approach to have points for changing (improving) the training approach.

The evaluation profile, see Figure 2, provides a basis for defining strengths and weaknesses. The elements of a training approach that does not fulfill the base practices (do not score as fully satisfied or partly satisfied) are than denoted as gaps that might be a risk for successful implementation. However, a more detailed and illustrative description is needed. The report that will be provided to a customer should therefore contain the evaluation profile, plus a list with positive and negative comments, a list with strengths and weaknesses and a list with improvement suggestions.

The first list presents positive and negative comments that were derived from the interviews about the training approach per phase, e.g., for the wrap-up phase. The list was meant to provide the customer with the motivation of ranking the APO training

approach as presented in the evaluation profile. Positive comments are the remarks of interviewed people that refer to those parts of the training approach that were performed conform to the base practices. Negative comments concern those that were not conformant. The comments were as much as possible written in the terminology specific for the training approach to present illustrative examples for the customer.

The second list presented the strengths and weaknesses of the training approach. The weaknesses address the risks of successful implementation of the training approach. The list was constructed by selecting the positive and negative comments (that already refer to particular base practices) and relating these to the levels distinguished in the Evaluation Object Architecture, see Figure 1. For example, the lack of author instructions for participants was caused by a deficit in the implementation of the training. Unclear responsibilities were caused by insufficient definition of roles. And insufficient quality assurance might be due to lack of attention for such assurance by the training approach. Such a list points out the origins of the detected weaknesses in the training approach. This is a more effective way in convincing the customer of the evaluation results than presenting only a list with strengths and only weaknesses. In total 8 strengths and 19 weaknesses in the training approach were defined. Next table provides the numbers of strengths and weaknesses reported according to the levels of the Evaluation Object Architecture.

Level in Evaluation Object Architecture.	Number of strengths	Number Of weaknesses
Specific for the training	1	4
Specific for a role	3	3
Specific to this organizer/implementation	2	4
Specific for the training approach	3	8

Table 2   Number of strengths and weaknesses reported according to the Evaluation Object Architecture

The third list produced for the reporting to the customer was a list with improvement suggestions. Nine improvement suggestions were defined, covering most of the 19 detected weaknesses.

The report sent to the customer contained an executive summary, an introduction describing the how of the evaluation and who was interviewed, the evaluation profile, the three lists discussed above and conclusions.

4   Implications for SPiCE trainings

In this section we discuss the application of our method for process evaluation to evaluate different providers of SPiCE trainings. We do not come up with a comparison of different SPiCE trainings, but rather provide a means to compare these providers.

   A typical SPiCE assessor training consists of two phases: the provisional SPiCE - assessor training and the competent SPiCE assessor training. The first training aims at providing industry professionals – often with a background in IT – the knowledge

and skills that is needed to gain the skills for performing software process assessments according to ISO/IEC 15504. These trainings are grouped events, provided in a setting of a four or five day course. As a result of these trainings the attendees should be able to navigate through the ISO/IEC 15504 standard, to understand and explain the use of process assessments, organize the resources (people) and select the means for the assessment, conduct an assessment (with guidance of a experiences assessor) and present the results of the assessment to the organization that is assessed.

The competent SPiCE assessor training aims generally at generating an assessment plan (select the relevant base practices), additional skills (like document reviewing) and learn how to lead assessments. Between the provisional and competent SPiCE - assessor training people have to practice. A competent SPiCE assessor trainee has to practice for 200 hours at least.

Evaluation object – SPiCE trainings are process-oriented trainings for some part, especially the experience that is needed to start with the Competent SPICE assessor training. The trainings (ranging from 3 to 5 days) are a direct face-to-face form of teaching.

Base practices (for training) – we think the base practices in planning, preparation and wrap-up phases, as we presented in Table 1, can be helpful to examine a SPiCE training organization. Because the SPiCE -assessors train themselves during their assessments, the execution is more dependent on the individual SPiCE assessor trainees. This also stresses the weaknesses of the SPiCE training concept: there is hardly any tracing on this part of the process only than the lead assessor.

Information sources – in SPiCE trainings only the role of the participant (the person that is trained to be a SPiCE -assessor) can be explicitly denoted. The SPiCE -trainer is usually only available during these training days. A process advisor/mentor and an expert advisor are not explicitly denoted in the trainings. Although we have observed that SPiCE assessors will regularly communicate with colleagues about how to conduct assessments. We even think that a lot of skills and competence of SPiCE - assessors is communicated along this line.

Evaluation instruments – we advocate use a questionnaire approach to evaluate SPiCE training success. Some providers apply this already. But SPiCE certification can apply these means to have a better overview on the quality of the learning.

5   Conclusion and outstanding challenges

This paper has argued that trainings on-the-job – that are inspired by the constructivist view of learning – require new approaches for their evaluation. The presented process-oriented evaluation method using base practices – inspired by the ISO 15504 of doing software process assessments – provided us such a new approach for evaluation. A set of strengths and weaknesses about the training approach is provided. The

weaknesses facilitate the identification of risks for which improvement areas can be defined to prevent future problems with the training approach.

The case study that reports on the application of the method for on-the-job trainings at a German telecom company shows that the method is applicable and that it is worthwhile to invest in further developing the method. A major challenge is to continue the *development of processes and base practices*. Although well-founded, only a preliminary set of base practices was available during the case study. Changing the processes and base practices was already required during the study. It is expected that reapplying the base practices – outside or within the evaluation of the same training approach – will improve them further. This can be done similarly as it is done during the development of the now world-wide accepted ISO 15504 standard. About ninety "SPiCE trials" were conducted to test out the design decisions of the core part of this ISO-standard (codename: SPiCE) and find out guidance to conduct the SPiCE - assessments as well as on guidance on the construction, selection and use of assessment instruments and tools (SPiCE, 1998).

We think that process-oriented evaluation will be applicable as well as useful for other types of on-the-job trainings than the trainings we have looked at in our case; see section 2. We have therefore suggested in this paper also to discus the application of a similar approach to evaluate training of SPiCE assessors, which we see as a type of on-the-job training.

Additionally, the evaluation method might benefit from developing *maturity levels* for training processes, like it is done for software process assessments (El Eman et al, 1998; Paulk, 1995). Applying such maturity levels to evaluate training approaches will help to indicate whether or not an implemented training approach has reached a particular capability, e.g., that the training processes are performed (purposes of the training are generally achieved) or even of a higher level, namely that the process is predictable (the process is defined and is performed consistently, within defined control limits to achieve goals; there is quantitative understanding of the process). The findings of educational research with regard to effectiveness of educational learning might form a useful source for the determination of maturity levels, e.g., to find out how many maturity levels are relevant for educational processes.

Acknowledgement

References

El Eman, K., Drouin, J., Melo, W. (eds) (1998*), SPiCE – the theory and practice of software process improvement and capability determination*, Los Alamitos, IEEE Computer Society, 486 pp.

ISO/IEC TR 15504-5 (1998), *Information technology – Software process assessment – Part 5: An assessment model and indicator guideline*, Geneva, 1998.

Kirschner, P., *The inevitable duality of education: cooperative higher education*, inaugural address, Maastricht University, 2000.

Gerstenmaier, J. & Mandl, H. (2002). *Constructivism in cognitive psychology*, in N. Smelser & P. Baltes (Eds.), International Encyclopedia of the Social and Behavioral Sciences (pp. 2654-2659). Oxford, Elsevier.

Paulk, M., (1995) The Capability Maturity Model – Guidelines for Improving the *Software Process*, CMU/SEI, Reading, Addison-Wesley.

Simons, P.R.J., van der Linden, J., Duffy, T. (2000), *New learning: three ways to learn in a new balance*. In P.R.J. Simons, J. van der Linden, & T. Duffy, (Eds.). New learning. Dordrecht: Kluwer Academic Publishers.

Van Petegem, W.E.A. (2000), *Virtual Business Teams: A new approach in engineering education*. In Michel, J. (Ed.), "The Many Facets of International Education of Engineers", Balkema, Rotterdam.