



Univation

Institute for Evaluation
Dr. Beywl & Associates GmbH
Hohenstaufenring 63
50674 Cologne
Germany
www.univation.org



Evaluation report of the Leonardo da Vinci- Project Peer Review in Initial Vocational Education and Training

August 2007

Sandra Speer

1. Introduction

Peer Review is a form of external evaluation with the aim to support the reviewed educational institution in its efforts for quality assurance and quality development. Peer Reviews have a long tradition within the higher education sector. In VET provision, experience with peer review has so far been marginal in Europe. In the Leonardo da Vinci project “Peer Review in initial VET”¹ a European standard procedure has been developed for voluntary use by VET institutions. It has a formative, development-oriented function. During two to three days a group of four peers is invited to assess the quality of different fields of the institution – chosen from fourteen predefined quality areas. The so called peers are an external group of experts, who work in a similar environment and have specific professional expertise and knowledge of the evaluated subject. They are “persons of equal standing” who are of the same rank as the persons whose performance is reviewed. The main product of the peer review project is a European peer review manual for providers of initial VET. The draft version of the manual has been tested within 15 pilot peer reviews in 8 different European countries between April 2006 and November 2006. The peer review model described in the manual can be customized in the extent and focus according to the needs of the VET providers. For the preparation of the peers additionally an online peer training programme has been developed.

The peer review pilot project includes an internal formative evaluation, which has been carried out by Univation, Institute for Evaluation as “evaluation partner” within the Leonardo da Vinci partnership. The results of the evaluation are presented in this evaluation report (on hand) and have been presented beforehand within the partnership.²

¹ Leonardo da Vinci-Project AT/04/C/F/TH-82000, www.peer-review-education.net

² Speer (2006a), Speer (2007).

2. Evaluation Design

The evaluation was designed for formative purposes. The quality of the peer review procedure shall be enhanced respectively the quality of decisions about it. The evaluation aims at facilitating the review of the draft manual.

In a first step a meta-evaluation of the manual was carried out. Secondly the pilot peer reviews have been evaluated and recommendations for the further development of the peer review procedure and the manual have deduced. For the meta-evaluation of the manual evaluation standards have been used as a framework for reflection. The use of the “programme evaluation standards” has been reflected within a CEDEFOP research project. The suitability for the use in the field of Vocational and Education Training could be shown (Beywl/Speer 2004a) and the use of evaluation standards in Europe is widely accepted (Beywl/Speer 2004b).

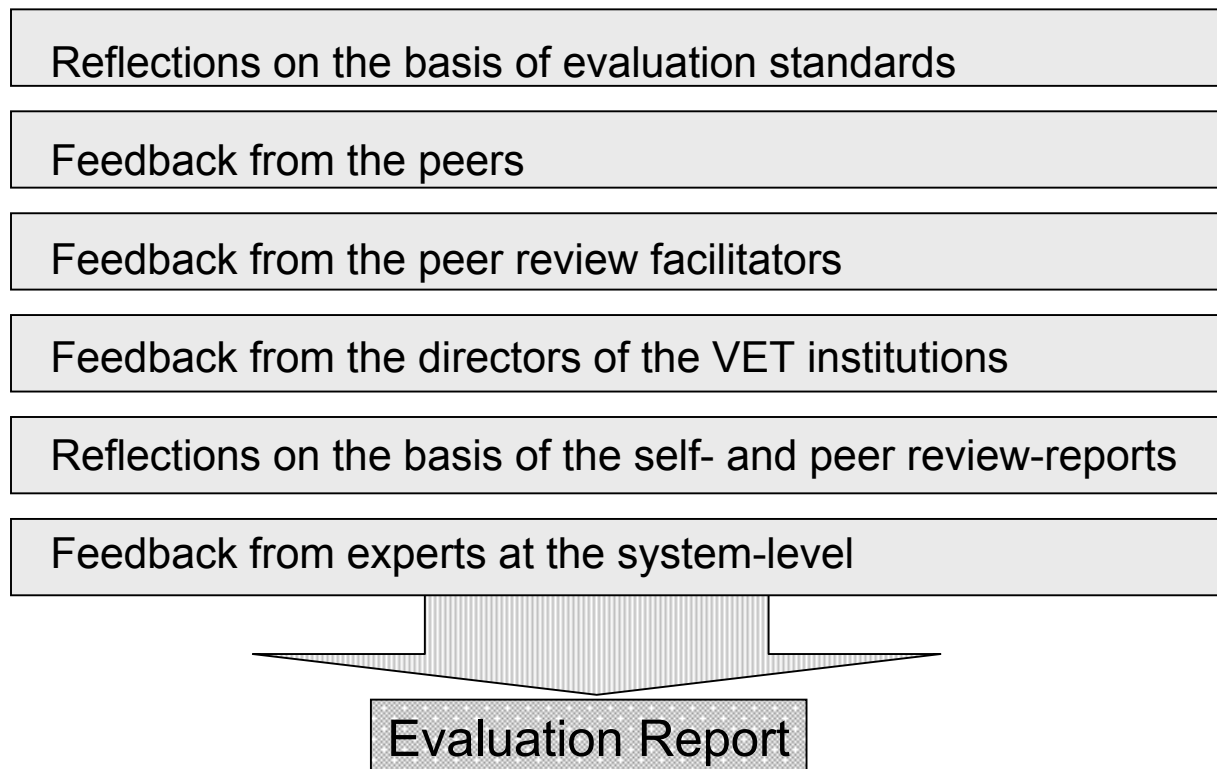
For the peer review pilots feedback has been collected from the main stakeholders – a triangulation of sources has been carried out.³ After each pilot the peers were asked to summarize their experiences with the peer review procedure as well as with the peer training programme and to write it down in the form “meta-evaluation”. Secondly each peer was asked to answer a questionnaire on the experiences with the peer review (“experiences peers”). Thirdly the peers were asked to bring in their feedback with a focus (more) on the further development of the procedure – for the first five pilots by telephone interviews and the following in written form (“questionnaire peers”). The peer review facilitators, who organized the peer review within the VET institution, have also been asked for their point of view of improvement (“questionnaire for OP reports”, “questionnaire facilitators”). Feedback concerning the peer review and benefits of the pilots has been given by the directors

³ See survey instruments in the annex. The survey instruments “telephone interviews peers”, “questionnaire peers”, “questionnaire facilitators”, “questionnaire directors” have been created by the evaluator in accordance with the participants of the transnational meeting in Trento. The forms “meta-evaluation” and “experiences peers” and “questionnaire for OP reports” were designed and sent out by the project management. In this evaluation all the answers to these different questionnaires – sent out by the evaluator and the project management - have been analysed in order to give a complete picture of the different pilots. The self-reports and the peer review reports were mainly based on the suggestions describe within the manual.

of the VET providers (“questionnaire directors”). The mostly qualitative answers in the questionnaires and the telephone interviews have been analyzed with the TextSortingTechnique (TST), which means that the text has been grouped around dominant issues.⁴ Within the following analysis aspects, which have been indicated many times by the respondents within the various questionnaires but also single statements, will be reported.

Additionally the “self-reports” of the VET providers, which they had to deliver before the peer review and the peer review reports, which were written by the peers as feedback for the VET providers under review have been analyzed, so that areas could be identified, where the providers and the peers needed additional information and where the manual might have to be adapted.

The following perspectives are included in the evaluation:



⁴ Therefore sometimes the reader will find abbreviations for the sources used within the part 4 of this report. The abbreviations are chosen in a way, that only the evaluator can identify the persons or providers and that they stay anonymously for the readers.

After the completion of the pilot phase experts on quality management in initial VET in the participating countries have been interviewed by telephone. They were asked to comment the manual and to give advice for the further dissemination and implementation

All the reflections on the peer review procedure are of relevance for the end product of the project – the manual – as the procedures are herein explained in detail and shall guide future peer reviews.


3. Reflections based on Evaluation Standards⁵

Peer review is an evaluation method itself, incorporating elements of self-evaluation, internal evaluation, and external evaluation.

For a critical look at the peer review method in initial VET four different evaluation standard sets have been used for meta-evaluation. The peer review method is an evaluation model itself, it is based on prior self-reports, sometimes personnel or personnel evaluation systems can be included into the peer review, and peer reviews can also include the evaluation of student evaluations or student evaluation systems.

The peer review manual has been reflected upon the following sources:

Sources for reflection



Peer Reviews as an evaluation model	Standards for Evaluation (DeGEval,2002) The JC Program Evaluation Standards (1994)
Peer Review based on prior self-evaluations	„Recommendations for the use of the German evaluation standards in the context of self-evaluation“ (DeGEval, 2004)
Peer Reviews evaluating personnel/ Personnel evaluation systems	The JC Personnel Evaluation Standards (2006)
Peer Reviews evaluating student evaluation systems	The JC Student Evaluation Standards (2003)

All these different standard sets quoted above are organized around the four important attributes of sound evaluation practice:

- propriety standards, which are intended to facilitate protection of the rights of individuals affected by an evaluation

⁵ This part of the evaluation report has been presented and discussed with the Leonardo da Vinci partnership at the transnational meeting in Trento, Speer (2006a);Speer (2006b)

- <u>utility standards</u> ,	which are intended to guide evaluations, so that they will be informative, timely, and influential
- <u>feasibility standards</u> ,	which are intended to recognize that evaluations are conducted in a dynamic, real-world setting and can be affected by many environmental factors
- <u>accuracy standards</u> ,	which are intended to determine whether an evaluation has produced sound information.

The evaluation standards from the four different standard sets have been used to guide the reflections of the evaluator on the peer review manual within this part of the evaluation.⁶ As all these standards are *maximum standards* it will not be possible – e.g. for any peer review handbook - to fulfill them all equally at the same time. Between some of the single standards trade-offs exist. However these standard sets can be used to reflect and improve current practices.

Four evaluation standards of the program evaluation standards (JC 1994) have been identified as especially important for the peer review procedure as laid down in the manual. These are “Information Scope and Selection” (U4), “Transparency of the Values” (U5), “Complete and Fair Investigation” (P3), “Unbiased Conduct and Reporting” (P4). So here the number of the peers and the composition of the peer team as well as the interaction with the VET institution and the quality of the self-report will be crucial and focused on within the survey described hereunder.

Based on the self-evaluation standards (DeGEval 2004) some issues seem to be central and also to be important for the continuative evaluation: the “context” of the VET provider should be described in detail (“Context Analysis”, A2). Furthermore it should be mentioned whether the self-report is based on a voluntary or compulsory self-evaluation and the underlying evaluation questions should be described in detail (“Described Purposes and Procedures”, A3) as well as possible different opinions between various stakeholder groups (“Transparency of values”, U5). The sources of information used in the self-report should be described in (enough) detail, so that the reliability and adequacy of the information can be assessed (“Disclosure of Information Sources”, A4) and it should be assured that no critical information is given to the peers (“Protection of Individual Rights”, P2). These aspects can be

⁶ These are reflections by the evaluator only and not on the basis of other stakeholders.

underlined in a future version of the manual respectively the self-report form and has also been focused on within the analysis of the self-reports as well as the survey.

On the basis of the evaluation standards of personnel (JC 2006), there are two aspects, which could be more highlighted within the peer review manual. First the qualification of the peers (“Evaluator Qualifications”, U3) regarding principles for sound personnel evaluation, which is sometimes relevant in the case of classroom observations, should be underlined. And the knowledge of the peers concerning national respectively regional VET systems and especially their compulsory evaluation systems could be included to the preparation of a peer visit. Secondly in case of observations the “validity orientation” (A1) should be stressed: for the validity of the peer review. It is important, whether observations are scheduled or unscheduled, whether the persons under review know the evaluation questions or not and who the classes or teachers under review selects.

Derived from the student evaluation standards (JC 2002), the following comments can be made for the manual: within the “Quality Area 3 Assessment and test” it could be stressed that the evaluation of students should provide information that identifies strength and weaknesses (“Balanced Evaluation”, P6) and that the student evaluation procedures should be practical, efficient and nondisruptive (“Practical Orientation”, F1).

The areas identified within these reflections in the light of the evaluations standards have also been considered in the following survey.

4. Results from the 15 Pilot Peer Reviews⁷

Fifteen pilot peer reviews were carried out within the Leonardo da Vinci partnership. The VET providers have been involved in the project from the beginning as so called “operative partners”. The VET providers are located in eight different European countries with very different evaluation cultures and represent various forms of organizations, such as public and private, small specialized and large multifaceted providers.

The following part of the evaluation report is based on various sources (see above, evaluation design).⁸ The response rates were as indicated in this chart:

Sources⁹	cases	target sample¹⁰	response	response rates in %
a) Meta-evaluations	15	15	8	53%
b) Experiences peers	63	63	34	54%
c) Telephone interviews peers	23	20	14	70%
d) Questionnaires peers	34	29	18 ¹¹	62%
e) Questionnaires facilitators	15	13	6	46%
f) Questionnaires for OP reports	15	15	3	20%
g) Questionnaires directors	15	15	5	30%
h) Self-reports	15	15	15	100%
i) Peer Review reports	15	15	9	60%

⁷ This part of the evaluation report has been presented and discussed with the Leonardo da Vinci partnership at the transnational meeting in Vienna, Speer (2007).

⁸ The 15 pilots were steered on the basis of a joint database of the Finnish National Board of Education and the öibf. The documentation of the 15 cases has been organised by the project management with the help of an electronic platform, which could be accessed by all the project partners.

⁹ See above; a), b), and f) have been distributed by the project management, c), d), e), g) carried out by the evaluator; h) has been prepared by the providers and i) has been written by the peers; h) and i) have been based on the procedures of the peer review manual. The evaluation is based on all the information from the different sources available until the 9th of February 2007. Therefore not only the instruments developed by the evaluator but also the other questionnaires and documents have been analysed to give a complete picture of the experiences from the pilot projects.

¹⁰ The difference of 10 cases between the total number of cases (column 2) and the targeted number of interviews (column 3) is resulting from the fact that some peers participated in more than one pilot. So it was possible to comprehend the experiences from all their pilots by one interview.

¹¹ Some peers filled out one questionnaire but were reflecting on more than one pilot they participated in. This fact has lowered the response rate to 62%; the experience from some more peers and pilots are reflected in these 18 questionnaires.

Each pilot peer review was a single case study. For some pilot peer reviews there were more forms “experiences peers” filled out and for others more “questionnaires”, but for each pilot peer review the picture received, was composed by various stakeholder feedbacks, so that it generated a credible basis – although the feedback from the directors could have been better.¹² The directors might have been the most difficult target group for the questionnaires as their workload is usually very high and sometimes their direct involvement in the peer review pilots was not so high. Secondly one director indicated that no feedback would be given until the final peer review report was sent. Some peer review reports have not been finalized and been sent out until the beginning of February.

This chapter is divided into three sections following the chronological order of the peer review process.

4.1 Getting started for the Peer Reviews

According to the self-reports, which were written by the VET institutions as a starting point for the peer review process, the reasons for undergoing the peer reviews were manifold. The most quoted reason for participating in the peer review was the formative aspect of learning from the peer team and to initiate further development, e.g. “we want to reflect upon our activities in order to improve our educational supply”. For some VET institutions the public relations aspect was important, e.g. “to help us make our efforts [...] transparent throughout the organization...”, “taking part in this review and getting the European certificate as first institution [...] can be very important for our public relations”. Other VET institutions wanted to check whether the peer review is a quality management tool for the future of the school or wanted to learn about VET quality management in general. One VET institution also mentioned the possibilities for international cooperation and another mentioned the possibility for European benchmarking efforts. So most of the providers were motivated by a formative evaluation approach of the peer review and in two cases the summative

¹² See annex.

aspect of getting a certificate and doing public relations with the participation in the peer review process dominated.

The participating peers indicated that they were motivated by testing and (developing further) enhancing the peer review process, learning about new quality management instruments, learning for the improvement of their own organisation and getting an inside view of a certain provider.

The transnational peer is essential for the international perspective of the peer review procedure. In many cases the transnational peer seemed to be the most inspiring input opening new horizons. The transnational peer can ask questions about topics, the national peers would perhaps take for granted. Furthermore the transnational peer is "allowed" to ask questions, which the nationals would not ask that directly, because it may seem to the nationals not to be politically correct. In other cases the transnational peer was the weakest part of the peer team and at the same time much has to be translated for the transnational peer. One peer wrote: „I am not convinced of the added value from an international peer review. The national peer reviews are easier since there is no language barrier. The international peer leans too much on explanations of his colleague peers.“ (QP§1#8).

One of the main difficulties of carrying out a transnational peer review is the language barrier. In three cases the transnational peer could speak the national language. But in many cases - especially in Nordic countries - the interviewed stakeholders had a very good command of English. In the cases, where the English language was relatively weak, the language often switched during the peer visit. The transnational peer was speaking in English and stakeholders were answering in their national language. The feedback session and its presentation was started in English, but in the moment the discussion started the national language was chosen. Therefore in these cases the transnational peer was partly relying on translation. Some VET providers organised a translator and sometimes a national peer could help translating.

The peer coordinator was the person coordinating the peer team. In many cases the coordinator organised the interview questions before the peer visit. This happened partly by e-mail partly within the peer meeting on the day before the peer visit. As stated by many peers, the process of collecting review questions beforehand is crucial. So in the peer meeting the already collected questions could be discussed, revised and grouped to the different interviews and interview persons. Secondly some peer coordinators organised a meeting beforehand with the VET provider. Also here the main objective was to find out on the priorities of the peer review and to organise the timing and the stakeholders to be involved. The peer coordinators found it very important and would mostly recommend organising such a personal meeting some time before the peer review. Additionally, It has been stressed by several peers that during the peer visit the peer coordinator should be a good (time)_manager as „some peers are not so much used to follow the procedure“ (IN§3#14) and a good communicator.

The peer teams have been composed very differently. As reported by the peers in one case the national peers were all professional assessors and in another case two „real“ peers – stemming from similar VET providers and having the same vocational background – participated. Most of the peer teams were completely mixed in their composition. One peer suggested that ideally at least one peer should have already participated in a peer review, but better even two peers should have experience with the procedure – which could perhaps be realised in the future, but not within the pilot project. “It’s important that not all the team members are administrators, each team should have at least one teacher” (QP§1#6). Another comment highlights that it might be better to have peers not from the same town than the location of the provider; because they get often interrupted during the peer review (MT#3). Additionally one peer suggested that “the international mix should be more explicit.” (QP§1#7).

Another suggestion of the peers was not to switch tandems¹³ during the process, as in many cases the peers did not know each other before. And one peer being a teacher in each tandem would be ideal (MT#3).

¹³ The peer teams of four peers were divided into two tandems working together.

The size of the peer team consisting of four peers has not been questioned: “The size of the peer team, four members, is functional. Four people are able to get a versatile view of the institute, but, on the other hand, it’s easy to reach consensus on observations thanks to the small enough size of the group” (QP§1#6).

For all the peer teams a very good cooperation within the peer team was reported. Only in one case conflicts occurred, which then could be solved to everyone’s satisfaction.

Within the project a web-based training was created and all peers were invited to use it.¹⁴ Out of fifty participating peers ten peers did not use it as a means of preparing the peer visit. One main reason of the peers not using it was that some of them were very well acquainted with the content, as they have been involved in writing the peer review manual and they felt well prepared. Persons using the web-based training were invited to post comments and examples of exercises they accomplished. But only 12 persons posted responses to all exercises of the web-based training. Though it can be concluded that the involvement on the discussion board was not very high, the peers, who did the exercises pointed out that the preparation in their native language turned out to be very helpful. Peers reported that the most important part of the web-based training was about „asking questions“ as it is different from the manual and the core activity during the peer visit. The users of the web-based training suggested for its improvement to include more animation, more exercises including control elements and to have the focus more on practical situations and examples. It was also suggested by the peers to extend this part of “asking questions” and to extent the training in general.

Many peers felt „well“ or „partly well“ prepared. They just felt „partly well“ prepared because either for (nearly) every peer it was the first time using the process and some indicated that they would have a need for further training. Answers to the

¹⁴ The members of the peer teams and the peer facilitators as well as the directors and other interested persons received a pass word for the web-based training, which was managed on a platform of the Aberdeen College and could be accessed via the project homepage www.peer-review-education.net.

question “How much time did you spend for preparing the Peer Visit?” (QF) varied very much. Some needed more than two weeks for preparing a peer visit, most of the peers between two and four days, and some peers one day or less. However, it has to be taken into account that the peers did not have previous experience and they had to learn how to implement the procedures. So in many cases the preparation time included an intensive study of the peer review procedures and not just only the preparation of the single pilot and its background information. Some of the preparation could be done during the travels, so that the time for preparation and travelling time overlapped. Especially the transnational peers needed more time for preparation as they sometimes also intensively informed themselves about the VET systems under review.

Single peers reported that they would recommend new peers to participate first as observers (QP§2#2). Ideally one international meeting per year would be organised for the peers (QP§2#10), so that an exchange of experiences and a further development of the process can take place. One peer suggested “a guided pre-peer review as training for both the staff of the VET institution, the students involved and the peer team - so-called “training on the job” (QP§2#10).

Each peer review pilot was organised by a peer review facilitator – a person from the VET provider under review, who should know the peer review procedure and be a contact person. In some cases the peer facilitators were also peers in other schools, so that they already knew the procedure. Some facilitators as well as some peers suggested to prepare checklists, which could help organising the whole process including e.g. informing staff on the project or the facilities needed (MT#3). Secondly the peer facilitators should be available during the whole peer visit, which is especially important for the transnational peer in case questions occur. A special training for peer facilitators was suggested by single peers and facilitators.

Most of the peers did not meet the VET provider before the peer visit (9 providers), whereas most of the Peer Teams organised a meeting for internal preparation (9 providers). In Finland and in Italy preparatory meetings were organised countrywide,

where representatives of the participating VET providers and peers were invited – but not all peers could participate.

The day before the peer visit the peers organised a peer meeting. In many of these peer meetings the following topics were on the agenda: getting to know each other (QP§2#6), the formation of peer tandems (QP§2#6), the definition of the questions for the interviews (QP§2#6) or the revision of the questions (QP§2#10), discussions with the management board of the VET institution (QP§2#10) and discussions of time limitations (QP§1#1). But also topics like “competence based education” (QP§1#8) or sharing the ideas upon the quality of the educational innovation (QP§1#7) turned out to be important for the preparation of the peer visit. So two single cases also general educational topics have been discussed at the peer meetings.

For the preparation of the peer visit each VET provider was asked to undergo a self-evaluation and to write a self-report as described within the peer review manual. This self-report was used by the peers to get an idea of the VET provider and its strength and weaknesses. According to the quality areas chosen, first review questions could be formulated by the peers. Although it was requested to send the self-report two months before the peer visit, most of them have been distributed less than one month before the peer visit. In many cases the self-report was received by the peer within the two weeks before the peer visit. Without the self-report in time it was difficult for the peers to prepare the review questions. The peer review manual says that the self-report should be send out one month before the peer visit. As the timing is critical for preparation of the peer visit one peer suggested to insist on the deadline for the delivery of the report and in case the self-report has not been delivered, the visit should be postponed (QP§2#10).

For most of the peers the self-reports turned out to be a valid basis for the peer review. The self-reports varied a lot in terms of quality and quantity. According to varying self-evaluation practices and knowledge about it, the self-report contained very different information basis. Sometimes the self-reports did not match the quality areas under review. Some self-reports consisted of more general information on the

VET provider and its educational program and were less specific about evaluative information collected beforehand. Sometimes a long annex was included, but single peers said that it is often not being read, even not printed out. In case that quantitative data is delivered, two peers said that it would be of help to get additional descriptive information, as the interpretation of numbers is difficult without knowledge of the context. In some cases the self-report did not include all information needed beforehand. The „description of the self-evaluation“ was sometimes a mix of describing self-evaluation practices in general within the VET provider and the self-evaluation process as a basis for this concrete self-report.

4.2 Carrying out the Peer Review

Most peers pointed out that the peer review was well organised in advance and on site. The timing of the peer visit was mostly very dense and some peers said that too little time was foreseen for reflection and preparation of the feedback. In one peer visit of two days the peer visit included 4, 5 hours for reflection and in another peer visit of two and a half days 7, 5 hours for reflection were included. The latter peer visit enabled the peers not only to prepare the feedback better, but also to work on the draft of the peer review report. Especially the latter example was recommended by the peers of that pilot for organising future peer reviews.

Most of the peers describe a good cooperation with the provider and just a few peers felt influenced. In the one case a peer felt under social desirability pressure it was mainly concerning sampling and the freedom to operate outside the providers indication (QP§1#10). In another case, some misunderstandings occurred, but they could be solved.

When the peers were asked about the number of quality areas under review, they answered according to their pilot experiences e.g.: “In my opinion there were too many number of quality areas [*five*]. We concentrated on the quality areas 1, 2 and 3 and we paid less attention to areas 5 and 7 (EX§2#16)”. “Two days for four Quality areas is too short”. (EX§1#13), “We did manage to get a good impression of the three quality areas”. (EX§2#14), “I think that a good review, in two days, should be based

maximum on two quality areas.” (EX§2#5), “I think that in VET providers with only a limited experience of this type of process, then one quality area in two days per tandem of Peers is about as much as can be effectively.” (EX§1#5). Summarizing it can be concluded that one, or two, or three quality areas were indicated to be feasible for a peer review of two days, but more seems to be difficult.

It could be shown that not always all criteria and indicators of a selected quality area have to be used for the review, but chosen aspects according to guiding review questions worked well and it turned out to be very useful for the VET providers – as it was suggested in the peer review manual. See annexe for an overview of the review questions suggested by the VET providers within six different quality areas. The quality areas are quite exhaustive and the system can become very heavy (QD§3#1). Not all criteria work well in every national VET context, e.g. the final exam-system is used in Finland as their assessment system is based on continuous assessment.

When asked, which quality areas were especially difficult to review, the following six were mentioned and reasons specified:

Quality Areas, which were especially difficult to review
QA 1 Learning and Teaching: was quite complex because of the high number of criteria and indicators (QP§1#10), (QP§2#6)
QA 3 Assessment and tests: because many laws and rules had to be known (from the perspective of a transnational peer) (IN§3#1)
QA 5 Infrastructure and financial resources: We had too less information upon these topics (self evaluation) and during the review it was difficult to gather the information. (QP§1#7)
QA 7 Institutional ethos and strategic planning: as QA5 (QP§1#7)
QA 11 Social participation and interactions: was difficult to realize, because many external partners should be involved and the timing of the peer review was not flexible (IN§5#11)
QA 13 Gender Mainstreaming: the underlying concept of „gender mainstreaming“ is not clear to all stakeholders, respectively may vary widely (IN§1#5)

The peers were asked about their experiences with the definition of the quality areas. The quality areas seem to be encompassing and VET institutions can be assessed in its complexity. Most of the peers like the wide range of indicators. However, as additional areas „citizenship“ and „environmental sustainability“ have been suggested

by one peer. One peer answered that a larger focus on results and the satisfaction of students on the various criteria instead of on plans and intentions of the provider would improve the peer reviews. Repeatedly it has been mentioned that the quality areas 1 and 2 have a strong relationship. In the interviews it was not always easy to separate them (QP§2#6). Different perspectives should be more clearly explained. Indicators can be same under different criteria, but the different perspectives should be seen and explained more explicitly (QP§1#6).

Concerning the sampling it was reported that in some pilots the number of interviews per stakeholder group seemed to be too limited. For reasons of practicability the providers have chosen the persons to be interviewed and organised the interviews (time and place) not the peers.

During the peer review many interviews and focus groups with teachers, students, company representatives, alumni, management, parents were included. But also other methods were used, like observations: tour of the school, classroom observations, company visits, photo-evaluation and analysis of documents.

In nine out of fifteen peer review visits classroom observations were included in the peer review design. Most of the interviewed peers had the impression that the teachers were informed about the classroom observations beforehand. It was assumed by the interviewed peers that not all of the teachers were informed about the quality areas chosen for the review. For the quantity of classroom observations not more than 4 classroom observations have been recommended by the peers. “otherwise the impressions get too manifold” (IN§1#14). One peer indicated that the choice between the observations of whole classroom sessions or „walk in walk out“ should depend on the relevant national evaluation culture. Additionally one peer suggested to integrate in the manual some examples for observation grids.

Concerning the preparation of the review questions and draft questions to be asked the feedback from many peers included the suggestion that the questions should ideally be collected by the peer coordinator in advance and the questions to be asked

should be based on the quality areas and the underlying indicators. Secondly the self-report and other documents of the providers served for the deduction of review questions. For the future one peers said that the identification of questions could be more shared among the peer team, e.g. division of questions for specific target groups or questions for quality areas. Some peers recommended asking overlapping questions - because of overlapping quality areas - at the beginning before going into specific questions. Some peers mentioned that a simple language for the questions to be asked should be used and that the formulation of questions is very important as especially students sometimes tended to answer what is expected from them (social desirability).

The staff of the VET providers reacted overall very positively. One director reported that the staff was sometimes very eager to take part in the visit. They appreciated discussing with peers instead of “audit authority”. The atmosphere was reported to be very open and reach of ideas. Another director mentioned that the staff was fully cooperative with the peer review team during their visit. “Full and frank exchanges were made in a fully professional way between true peers. All staff involved was pleased to be able to commit efforts to this useful process. Internal feedback suggested an interest in, and support for this process of review by critical friends” (QD§5#1).

4.3 Learning from the Peer Review

At the end of each peer visit different formats for the oral feedback have been realised - differing in the composition of the feedback audiences: principal and peer facilitator only, or senior management only or senior management and involved teachers or senior management and involved stakeholders or all staff members and involved stakeholders or staff and all students invited. Concerning actors of the feedback session often all the peer team was involved, sometimes just one peer gave feedback or two peers (one national and one transnational peer). In all except one case the feedback was given directly (live) and in one case the feedback was

registered and been shown on a later date. Nearly all the peers were satisfied with the way they provided feedback. Also in the case of the video feedback the facilitator reported that the staff reacted very positive and could learn from it. However, to learn more about this special case, the staff should be interviewed directly.

An analysis of the pilot peer reports by the evaluator showed that most of the reports were very concrete, described all the peer review processes and gave detailed feedback. Just two peer reports were not so concrete and included only a few examples illustrating the areas for progress. The data collection methods and sources (3.3) were sometimes difficult to understand. In some peer reports the sources were mentioned in the overall assessment of each Quality Area and in others not. In one report it was written that the assessment was limited, e.g. “generally this criterion is adequate, but it is not assessable and it seems to be rather unclear for documentation's lack” (QA11), which helps to interpret the results. In some reports the assessment of a criterion was like a mark „fairly satisfactory“ or „we appreciate this criterion with the mark excellent/ very good/ good“. The peers reported that the format for the peer review report as included in the manual contains too much repetition and redundancy (3 times) (EX§1#15) and could be improved. Many reports were finalised weeks after the peer visit.

When the peers were asked to reflect about their lessons learned from the peer reviews, many mentioned the reflection of their own competencies, e.g. pedagogical and intercultural competences and that they could improve and their own practices. Additionally many peers indicated that they could profit from this new assessment experience. One peer said, realising that other providers are doing not much better was also useful. Another peer could report, that examples from the peer review have already been used for staff development training. Most of the peers would like to be involved again and gave overwhelming comments. Just two peers did not answer the question within the questionnaire and one indicated that it is not within the professional priority. One peer would need some incentive like an additional higher remuneration or travel to a country not visited before for participating again in a peer review.

The VET providers being asked for their benefits they answered that their own goals for initiating the peer review have been reached. E.g. a positive effect was that the goals set by the headship have been shared by teachers and staff (QD§1#11). In some cases the implementation of (slight) changes has already started respectively plans and preliminary schedules for improvement areas have been developed. It could be learned from best practices from other schools and contact to schools abroad could be established. The positive effects reported by a provider included a constructive critical and independent external analysis of their business practices and a closer working relationship with their benchmarking partners.

One VET provider would recommend the peer review procedure to other VET providers „under the conditions that the school's management is willing to be open and honest in all matters and does not consider it a problem to get negative feedback!“ (QD§1#13). Another VET provider would just recommend it in collaboration with national review institutions (QD§1#8)¹⁵ and a third VET provider would recommend it under no preconditions (QD§1#11).

Asking the different stakeholders for feedback concerning the manual itself, most of the answers indicated that the content is very complete and well structured. One idea for improvement was to introduce short visualized charts of processes, which ought to be self explanatory and transparent to students and staff (QP§2#10). In general, more advice is needed for preparation of the different stakeholders and more detailed information in certain sections (see above).

¹⁵ Here it would be interesting to further investigate, why this VET provider recommends the collaboration with the national review institutions.

5. Feedback from national VET Experts

After the pilot phase has been finished from January until March 2007 the expert interviews were carried out. Most of the experts interviewed stressed how extensive and already encompassing the peer review manual is. In most of the European countries involved in the project peer review procedures are still missing in initial VET and the experts underlined the usefulness for the future. In the following summary of the interviews the focus is on suggestions for the further development of the manual and dissemination as well as implementation in different national contexts.

Austria

Austria is interested in actively promoting the Peer Review Process on the European level (working program 2008-2009). The peer review has already been reflected in an Austrian working group and it seems to be appropriate for the national VET system. From the Austrian perspective the infrastructure for the peer review can be either on national or on international level. The inclusion of an international peer would be appreciated in case the VET providers also agree to it.

Denmark

In Denmark a „benchlearning“ approach has already been implemented. This benchlearning has a more quantitative orientation. The benchlearning and the Peer Review could not be well combined as the first is not very open for other systems.

In the Leonardo Peer Review Manual a “fair standard” should be defined as the VET systems vary widely through Europe and it gets difficult respectively impossible to compare them. The training of the peers seems to be crucial - especially the knowledge of the transnational peer on the national VET system under review. For example in Denmark the national curriculum is predefined, so a review could just look how it is implemented in the relevant context.

It has been suggested to integrate „European standards“ into the Quality Areas in the Peer Review Manual and additionally it has been suggested to make QA14 compulsory.

The association of Danish VET providers could be perhaps a good network for further dissemination.

Hungary

Quality Management is obligatory for schools, but they are free to choose a model. The Peer Review Procedure could perhaps be integrated into the Quality Management of schools. However, the costs of the approaches are crucial for the choice of the providers.

No need for change was identified. In a first step information about the Peer Review has to be spread, ideally via the Hungarian National Institute for VET and Adult Learning. The homepage would be the easiest way to use for publication.

Italy

The Peer Review should be introduced in Italy on a voluntary basis. There should be a public agency organising the process. A current debate about a national agency for quality is ongoing. However both parts of the VET education system – schools and training centres – should be covered, so maybe two agencies should be involved. The resources of the schools are currently concentrated on accreditation. Ideally after the completion of the Peer Review Manual it should be disseminated via the Italian Ministry of Education to all schools. The introduction of the peer review would not be an automatic process and instruments to make people know the process and to understand the importance of this kind of procedure have to be developed. There should be a public institution, that helps the schools to find a partner, understanding the process, the manual, finding a peer and training the peer in the first phase.

Portugal

The peer review concept could be very useful for Portugal. The implementation should be voluntary as a compulsory implementation might be unable to reach the aims of the project. No need for change was identified. It was proposed to work on a

list of national and regional institutions, which could be interested in the peer review and then to stimulate their involvement in the implementation and in the validation of the results as well as meta-evaluation. For disseminating the results of the peer review project an information program on the objectives and methodologies to key persons in Portugal was proposed. The willingness for future collaboration has been indicated by the expert.

The Netherlands

The peer review concept could become widespread on a national level. In the Netherlands there exists already a high culture of self-evaluation. For the self-evaluations usually one external person (peer) is involved. Additionally external inspections take periodically place. So the involvement of peers is already realized in the VET evaluation system and the different perspectives (internal/peer/external) within the VET quality management systems are already partially linked.

For a transnational European peer review, the costs and benefits might be unequal. The costs for the VET provider are high as they have to prepare all the documents in English and have to pay for the international peer. The Dutch VET providers do not see the need for it. If it is subsidized in the future it might be more used. On the other hand the Dutch VET providers also want to learn from others and the best way seems to participate as a peer. So generally there is the interest of single persons for the future to carry out a peer review. The dissemination concerning the results of the peer review project takes already place through a Dutch network of Quality Assurance in VET.

Romania

Peer Evaluation is seen to be a useful instrument between the already existing self- and external evaluation. For initial VET inspection exists - quality inspection every year and also external evaluation - realised by the national agency for quality assurance.

Peer-evaluation could be a necessary and very useful tool in order to make the development and quality processes more permanent. There is the intention of encouraging Peer Review in schools and perhaps to link it with benchmarking.

The introduction of peer review will be discussed on a voluntary basis – mainly for cost reasons. More flexibility in the school budgets is a prerequisite for financing a peer review procedure in Romania.

The peer review manual was considered to be too normative and a review of the manual could help to make the approach more flexible and to show options for the head teacher to choose from. The external and peer evaluation could be linked – some of the external evaluation could be done by peers instead of inspectors. It was suggested to introduce beside the existing quality areas aspects on minorities as well as on citizenship.

The Romanian Agency for Quality Assurance could propose to discuss the peer review on the national level, however the National Centre would be responsible for it.

United Kingdom

The Quality Improvement Agency for Lifelong Learning (QIA) has been carrying out and supporting work called "peer-referencing", more or less like peer review as an assessment function, but also with an improvement focus. In the last five years, there has been a strong focus on external inspection methodologies, although self-assessment has been seen as a critical process. In the last year, confidence in self-assessment has grown, as capabilities in self-assessment have been demonstrated. So there is a corresponding shift towards peer-review methodologies.

At the moment peer-review is developed as a voluntary approach. The QIA is leading a lot of work on peer-review. It is currently 8 national pilots, where colleges are working together in groups to carry out peer review. The QIA is also introducing peer-review methodologies within the quality improvement programmes it supports. So the QIA is been set up to develop programmes to support quality improvement in the

sector and it wants to use peer review as a methodology in doing that. The government is now to announce proposals on self-regulation and that may mean, that the peer-review over time, perhaps in a couple of years, becomes developed as more of a national system, as part of a regulatory mechanism.

Probably the biggest barrier to extending this work to other providers to wider sectors would be cost and time, staff time. Other important issue is competition as many colleges of the pilots involved in the UK are working together in a twenty kilometres-radius. The inspection in the UK could move, if it wanted to, towards a system, where it almost validates peer judgement. It is quite possible that peer-review type work can be used not only to assess performance against criteria, but can be used to support the development of the criteria in themselves.

6. Summary

The peer review has been practiced very differently within the 15 pilots. Each peer review was a single case with its particular VET context, particular evaluation culture, different peer review teams and forms of peer review designs chosen within the peer review procedure as described within the peer review manual. This kind of flexibility is the basis for the peer review process. For nearly all of the peers it was a very positive experience. The peer reviews could give valuable feedback to the VET providers.

The peer review procedure - as explained in the manual - was feasible and very useful. The evaluation could identify, where differences between the pilot peer reviews and tensions within the various forms of the peer review processes chosen occurred, which will be described hereunder and guide the summary.

Simple vs. professional peer reviews: For some VET providers the peer review procedure was the starting point for a general quality improvement process in their institution and for other VET providers the peer review was a new tool in their already elaborated evaluation culture. The beginners were confronted with a high workload for preparing the self-reports, the peers had to work much on data collection fitting to the local context, but got worthwhile feedback to initiate urgent and effective improvements. Experienced VET providers use self-evaluations routinely in formative and summative ways, and therefore it is easy for them to introduce the peer review as a new method, they already have a sound data basis and knowledge in organizing similar reviews. It turned out to be more difficult for the peers to identify areas of improvement for very professionally managed VET providers.

Transnational vs. national peers: For some peer reviews the outside view of the transnational peer was a condition sine qua non for new experiences, insights and ideas. For others the lack of proficiency in the national language and/ or of knowledge concerning the local VET system turned out to be a crucial weakness of the transnational peer. In some cases the transnational peer seemed to be the

“weakest” peer review team member and in others the most inspiring one. Ideally it should be a win-win situation for peer and VET provider.

Many vs. few selected Quality Areas: Some pilot peer reviews were designed as a broad peer review with many quality areas, others focused on specific questions within a few selected quality areas. Only very experienced peers are able to review sufficiently enough more than three areas within two days. Most of the peers recommend to review two or three quality areas.

All vs. selected criteria: In some cases the pilot was a more standardised peer review using all the criteria of the quality areas selected for a general review by the providers. Questions to be asked by the peers have then been derived by the help of the criteria – and also sometimes the indicators mentioned in the peer review manual – as well as on the basis of the self-reports. Other peer reviews focussed responsively according to specific questions of the VET provider on selected indicators and criteria. The use of all criteria and indicators for certain quality areas would allow for a better comparability and possible future certification. On the other hand some VET providers are especially interested in the review of specific review questions and not in the review of all indicators.

Few vs. many review methods: Each of the pilot peer reviews consisted of many interviews - most of the time in form of focus groups. Some pilots used creatively additional review methods, e.g. photo-evaluation. In case that classroom observations were chosen – which are usually time-consuming – there was less time left for the interviews. Each review method has its advantages and disadvantages. For reasons of time and resources a choice of methods has to be made and this is in line with the manual.

Data collection vs. analysis: When as much time as possible was used for the review itself, little time was left for documentation and reflection among the peer team and vice versa. A well prepared feedback needs time for preparation and good documentation during the peer review in order to facilitate writing the peer review report, as the peers do not meet personally again after the review.

Independent vs. efficient sampling: For reasons of practicability the VET providers were organizing the interviews and determined the persons to be interviewed. This can lead to an unbalanced review in case that the provider does not look for a “realistic” mix of interviewees.

Feedback for a few vs. large group: Different kinds of feedback sessions led to different effects, e.g. some peers gave their feedback to the management only others to a big stakeholder group (e.g. 60 persons) other peers registered a feedback video. The feedback varied between diagnosis and organizational development.

The interviews with national experts in VET and quality management from eight countries showed that there is a high interest in the peer review procedure as described in the manual. Most of the experts interviewed stressed how extensive and already encompassing the peer review manual is. In most of the European countries involved in the project peer review procedures are still missing in initial VET and the experts underlined the usefulness for the future and want to take an active part in the further discussion and dissemination.

7. Recommendations

The main practices and tensions within the peer review procedure have been described within this evaluation report and highlighted within the summary. On the basis of these tensions, decisions could be taken by the project management to direct the peer review method in either direction of the tensions and to revise the manual accordingly. Alternatively it would be recommended to stress the existence of these tensions within the manual, so that the future users can easily understand their options of organising the peer review and know where to pay special attention. The rich experiences from the pilots can illustrate these options and show how the procedures have already become alive. Recommendations for single issues are explained in the following sections.

In countries, where English is not spoken fluently by most stakeholders, it would be very helpful that the transnational peer is fluent in the national language of the peer visit country. The VET providers would prefer to get the feedback in their national language. In case that translation is needed, a translator should be available for the entire time of the peer visit – the translation should not be done by one of the peers because they already have enough tasks to accomplish

The pilots showed that the transnational peer should be extremely qualified for being able to contribute actively to the peer review and should be willing to take more time for preparation than the national peers as more information on the national context of the VET system has to be studied beforehand. The transnational peer needs a very high cultural sensitivity especially for the feedback session. Also first national experience with the peer review procedures could be an interesting step before becoming a transnational peer. How to select the transnational peers will depend on the way the peers will be matched in the future after this pilot phase.

Potential additional and important topics for a peer training could be identified – from the interviews with the peers and the experience of the evaluator -, such as “awareness of different educational cultures and evaluation cultures”, “triangulation of

sources and methods”, “non-reactive review methods” as well as the extension of the topic “asking questions versus review questions”. These topics could be either integrated into the existing web-based training or become part of future training manual.

For the self-report of the VET providers it would be helpful to integrate two separated paragraphs on a) evaluation culture: evaluation in general within VET provider under review and b) process and methods underlying the relevant self-report. Secondly a detailed description of the VET context should be more encouraged. Both could be done with slight adoptions within the peer review manual. An example of best practice self-reports could be distributed to future VET providers under review or on the project homepage. Furthermore a system of approval for self-reports by the peer coordinator or by a central organising institution should be taken into account – depending on the future further development of the peer review initiative. The approval by the peer coordinator could be realised with less bureaucratic burden. On the other hand a central organising institution could better have the overview.

Concerning the planning and timing of the peer review visit, enough time for reflection within the peer team, for the preparation of the feedback session, as well as for writing a draft peer review report should be included. The time table in the peer review manual should be adapted.

Many peers mentioned that the quality areas should be revised. The revision could include the underlying educational concept. In general the comprehensibility of the criteria could be improved and especially overlapping areas explained or deleted. A differentiation into core criteria and peripheral criteria could be helpful. Some of the quality areas (1, 3, 5, 7, 11, and 13), which turned out to be especially difficult to review should be complemented with additional information – e.g. illustrations from the pilot reviews - in the peer review manual.

For the sampling within the peer review process it is crucial to include critical voices and to have a realistic mix of interviewees. During the pilots the interviews have been

prearranged by the providers, but a random selection of interview partners could in future also be included, where possible – e.g. for student interviews.

For the classroom and other observations grids turned out to be very helpful and examples for grids could be included into the peer review manual as the peer team should decide first what to observe for not being too overwhelmed by the many impressions they get during the visit and especially the observations.

The focus group method – which is explicitly mentioned in the peer review manual turned out to be a very efficient instrument for data collection. It can be used as an introductory data collection instrument within the peer visit and also for validation. Its use for “homogeneous” group interviews could be more clearly explained within the peer review manual and the “moderation technique” could be introduced as another useful method for data collection within “heterogeneous” groups – which are also already mentioned in the manual - as not every peer has previous experience with these two different data collection methods.¹⁶

The feedback sessions are a very important part of the peer review process and sometimes ice has to be broken, when the provider is not used to receiving „critical“ feedback from outside and is e.g. more used to quality management audits. During the pilots it turned out that the setting of the feedback can give the peer review a different notion - from diagnosis to organisational development. Various important aspects for arranging feedback sessions are explicitly explained in the manual. Their strengths and demands on the peers could also be mentioned at this point of the manual and further evaluation of this aspect is recommended by the evaluator.

For the peer review reports some more help could be included in the manual. Best practices for the level of detail and for the language of writing such a report could be mentioned. Additionally it should be stressed that it is recommended by the evaluator not to include marks in the report – as it is stressed in the manual, that the definition

¹⁶ See Peer Review Manual for Initial VET 2005, p. 31); the “moderation technique” is also known under the name “meta-plan-technique”.

of the quality of the VET providers depends on various factors,¹⁷ and the peer review is not a summative review based on defined European standards - and that the delivery of the report should be close to the time of the peer visit.

The pilot peer reviews and their evaluation has led to a wide array of documentation from peer review practices. From now on “real life situations” of the peer reviews could be introduced into the manual or alternatively for the future the development of an additional training manual would be recommended by the evaluator.

¹⁷ See European Peer Review Manual for initial VET 2005, p. 44 f.

Bibliography

Beywl, Wolfgang/ Speer, Sandra (2004a), Developing standards to evaluate vocational education and training programmes. In: Descy, Pascaline.; Tessaring, Manfred. (eds) The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report. Luxembourg: Office for Official Publications of the European Communities, 2004.

Beywl, Wolfgang/ Speer, Sandra (2004b) and Literature-Based Reflections on Western European Evaluation Standards and Practices. In: New Directions for Program Evaluation, no. 104, Winter 2004. 43-54.

DeGEval Deutsche Gesellschaft für Evaluation (2002), Standards for Evaluation, Köln.

DeGEval Deutsche Gesellschaft für Evaluation (2004), Recommendations for the use of the German evaluation standards in the context of self-evaluation, www.degeval.de.

European Peer Review Manual for initial VET, Vienna 2005.

Joint Committees on Standards for Educational Evaluation (1994), The Program Evaluation Standards, How to Assess evaluations of educational programs, Thousand Oaks.

Joint Committees on Standards for Educational Evaluation (2002), Student Evaluation Standards, How to Improve Evaluations of Students, Thousand Oaks.

Joint Committees on Standards for Educational Evaluation (2006), The Joint Committee Personnel Evaluation Standards, draft 2nd Edition.

Speer, Sandra (2006a), Evaluation within the Peer Review Project, Presentation held in Trento, Leonardo da Vinci Transnational Meeting, 17th of January 2006.

Speer, Sandra (2006b) Peer Review – Peer Evaluation, Presentation held at the European Evaluation Society (EES) conference in London, 05. October 2006. www.europeanevaluation.org

Speer, Sandra (2007), Evaluation Results, Leonardo da Vinci Transnational Meeting, Presentation held in Vienna, 12th of February 2007.

