
Transnational European Evaluation Project

Methodological Reflections

ENQA Occasional Papers 6

TEEP was financially supported by the European Commission



European Network for Quality Assurance
in Higher Education • Helsinki

Cover design: Jussi Hirvi / Green Spot Media Farm
Page layout: Anja Kauppi / Pikseri Julkaisupalvelut

© European Network for Quality Assurance in Higher Education, 2004, Helsinki
Copying allowed only with source reference.

ISBN 952-5539-00-8 (paperbound)
ISBN 952-5539-01-6 (pdf)
ISSN 1458-1051

Printed by Multiprint
Helsinki, Finland 2004

Foreword

The Transnational European Evaluation Project (TEEP) was a pilot project, which investigated the operational implications of European transnational quality evaluation of study programmes in three subject areas: History, Physics, and Veterinary Science. In total, fourteen programmes in eleven different European countries were evaluated.

The objectives of TEEP were to develop further a method for transnational external evaluation, building on experiences, such as the TUNING Project and the Dublin descriptors developed through the Joint Quality Initiative, to identify potential obstacles to transnational evaluation and to indicate strategies that may be used to overcome them, and to contribute to greater awareness, transparency and compatibility within European higher education. The overall conclusion of this report is that these objectives were indeed satisfactorily met.

One of the aims of TEEP was to test the use of common criteria. It is important to emphasise that this criteria approach has provided the basis for making comparisons possible. The common criteria have functioned as shared reference points, and ensured that the same topics were evaluated across the three disciplines and the 14 programmes. Any future transnational evaluation project should establish criteria/reference points that are compatible with national and local contexts, and use terminology that is, to a large extent, familiar and useful for the programmes being evaluated.

The TEEP project has made a significant contribution to stimulate discussions about and recognition of the need for the programmes to develop explicit quality assurance strategies. The project has also shown that when national states have committed themselves to political objectives (aligned to the Bologna process) it is easier to reach a common interpretation. In that respect the project has provided a valuable insight into the condition for the implementation of the Bologna process at programme level.

This methodological report is the final phase of the TEEP project. The publication of the TEEP methodological report and subject reports, available at www.enqa.net/pubs.lasso, contributes to greater transparency and increasing awareness of the potential and importance of compatible approaches to quality assurance within a transnational framework.

It is my hope that the reader will find this report useful and of value.

Christian Thune

Chairman

ENQA Steering Group

Contents

1	Introduction	5
1.1	The Bologna Declaration	5
1.1.1	<i>European transnational projects on quality in higher education</i>	5
1.2	Introduction to TEEP	7
1.2.1	<i>Anticipated benefits from TEEP</i>	7
1.2.2	<i>Report structure</i>	8
2	Executive summary	9
3	Results from the subject-specific reports	11
3.1	Degree structure and definition	11
3.2	Competences and learning outcomes	12
3.3	Quality assurance	13
4	Methodology	15
4.1	Organisation	15
4.1.1	<i>Process</i>	15
4.1.2	<i>Scope of evaluation</i>	16
4.1.3	<i>Recruitment and selection of programmes</i>	17
4.1.4	<i>Budget</i>	18
4.1.5	<i>Language</i>	18
4.1.6	<i>Selection and composition of the expert panels</i>	19
4.2	Self-evaluation process	20
4.2.1	<i>Self-evaluation manual</i>	20
4.2.2	<i>Launch meeting</i>	21
4.2.3	<i>Focused approach</i>	22
4.2.4	<i>Self-evaluation process</i>	24
4.2.5	<i>Self-evaluation group</i>	24
4.2.6	<i>Quality of self-evaluation reports</i>	25
4.3	Site visits	25
4.3.1	<i>Organisation of the site visit</i>	25
4.3.2	<i>Interview guides</i>	27
4.4	Report	27
5	Results and conclusion on the use of criteria	29
5.1	Criteria-based evaluation	29
5.1.1	<i>Formulation of the criteria</i>	29
5.1.2	<i>Applicability of the criteria</i>	30
5.1.3	<i>Conclusions</i>	31
6	Reflections on methodology	33
6.1	Reflections from the management group	33
6.2	Reflections from the institutions	35
6.2.1	<i>Report from the evaluated Physics programmes</i>	35
6.3	Reflections from the expert panels	36
6.3.1	<i>Comments on the TEEP Methodology from the Physics Expert Group</i>	36
7	Appendices	39
	Appendix A	39
	Appendix B	40

1 Introduction

1.1 The Bologna Declaration

Since 1999, European perspectives on the quality of higher education have been strongly influenced by the follow-up processes to the Bologna Declaration of that year, signed by 29 European Ministers of Education. By signing this declaration, the Ministers agreed on coordinating their policies towards achieving a number of objectives, which they considered to be of primary relevance in establishing a European area of higher education and promoting the European system of higher education worldwide.

This general background, together with subsequent initiatives and developments occurring between the ministerial meetings in Bologna, Prague and beyond, have provided the main motivation for setting up the Transnational European Evaluation Project (TEEP).

TEEP was supported by the European Commission through the SOCRATES programme. It was part of a package of measures initiated by the European Commission in order to stimulate the Bologna Process (*“From Prague to Berlin, the EU contribution”*). The project was coordinated by the European Network for Quality Assurance in Higher Education (ENQA), with the participation and contribution of the SOCRATES Thematic Networks of the three respective disciplines History, Physics and Veterinary Science. Representatives of ENQA, the chairpersons of the SOCRATES Thematic Networks, representatives of the European Commission and representatives of the relevant quality assurance agencies constitute the management group for the project.

1.1.1 European transnational projects on quality in higher education

There are a number of projects that are of particular relevance to the establishment and development of TEEP. The most important projects are:

- The wide-ranging European Pilot Projects conducted in 1994/1995, supported by the European Commission. Seventeen countries, the fifteen EU members as well as Norway and Iceland, were involved in this project in which a total number of no less than 46 programs within higher education were evaluated simultaneously. The main idea of the project was to test a common methodology for programme evaluations, which was at the same time suitable for national adaptations.
- The international evaluation of electrical engineering programmes in Belgium, the Netherlands, Switzerland, Sweden and Germany, initiated by the Dutch Quality Assurance Agency, VSNU, and conducted in 1991/1992. The purpose of this project was to reach a mutual understanding and recognition of diplomas of the chosen programmes across the countries involved.
- The international research project initiated by CHEPS (Center for Higher Education Policy Studies) and conducted by researchers from the Netherlands, Germany and the UK is another example of an international evaluation. In this project from 1991/1992 ten programmes in Economy from the three countries mentioned above were evaluated. The project was primarily oriented towards methodological development. More specifically the aim was to develop a valid, reliable and effective methodology for comparing educational quality across the systems of higher education in a number of European countries.

- The two-year project TUNING Educational Structures in Europe, launched in May 2001 and organised by European universities and supported by the European Commission through the SOCRATES programme (<http://www.relint.deusto.es/TUNINGProject/index.htm>). The project aimed within the context of the Bologna process to “tune” educational structures in Europe, open a debate on the nature and the importance of subject-specific and general competences involving stakeholders, identify subject-specific and general competences and lastly develop the use of ECTS credits. The first phase of the project came to an end in May 2002 and a second phase to test the key findings in the project has started since then.
 - The Cross-border Quality Assessment of Physics conducted in 2000/2001 that involved five programmes from four universities placed in three different countries. Four national/regional quality assurance agencies were involved in the conduction of the evaluation. The aim of the project was to compare the programmes and to analyse whether students received equivalent qualifications. The method applied for the evaluation drew heavily on the lessons learned from the evaluation of engineering programmes mentioned above. The overall approach with an international committee responsible for formulating minimum requirements and conducting the site visits resembled the one used in the evaluation of engineering programmes. However, the principles behind the composition of the international committee differed. In the Physics evaluation it was decided that the committee members should all be independent of the participating institutions.
 - The International Comparative Evaluation of Programmes in Agricultural Science conducted by the Danish Evaluation Institute (EVA). The evaluation included programmes offered in Denmark, Germany, Ireland and the Netherlands. The evaluation was a Danish reflection of the Bologna process and the specific objective of promoting European cooperation in quality assurance with a view to developing comparable criteria and methodologies.
- European ministers have recognized the vital role that quality assurance systems play in ensuring high quality standards and facilitating the comparability of qualifications throughout Europe. They have also encouraged closer cooperation between recognition and quality assurance networks and sought to promote European cooperation in quality assurance.
- Whilst debates continue about the relative roles and merits of different quality assurance approaches, several notable initiatives have been established. These include:
- Development of the roles of ENQA. Reflecting on the Bologna process, the EU Ministers of Education have assigned responsibility for the quality assurance development in higher education to the ENQA network. The ENQA network is supported by the European Commission through the SOCRATES programme. ENQA has taken actions to disseminate information, experiences, good practices and new developments in the field of quality assessment and quality assurance in higher education between interested parties, public authorities, higher education institutions and quality assurance agencies;
 - The pilot scheme ‘Promoting a “quality culture” in universities’ will help universities introduce internal quality assurance mechanisms that they can consider their own. The project is supported by the European Commission and conducted by the European University Association (EUA) (<http://www.eua.be>). The expected outcome is to create a critical mass of universities having concrete experience with internal quality assurance mechanisms helping them to improve their quality levels and being better prepared for external evaluations;
 - The Joint Quality Initiative supported by the Dutch and Flemish governments, and in particular the development of the Dublin descriptors, which are shared descriptors for first and second cycle degrees (<http://www.jointquality.org/>).

1.2 Introduction to TEEP

The Transnational European Evaluation Project (TEEP) was a pilot project with the objective of investigating the operational implications of a European transnational quality evaluation of study programmes in three subject areas: History, Physics and Veterinary Science.

The three subject areas of Physics, History and Veterinary Science were divided respectively between five, five and four participating European universities. In total, fourteen programmes in 11 different European countries were evaluated. For a list of the participating programmes, please see appendix A. In Physics and History the scope of the evaluation was confined to the first cycle degree level or equivalent. In Veterinary Science the full one cycle programme¹ was evaluated.

The objectives of TEEP were:

- To further develop a method for transnational external evaluation, building on experiences such as the TUNING Project and the Dublin descriptors developed through the Joint Quality Initiative, using common criteria on the basis of an evaluation process in three different disciplines.
- To identify potential obstacles to transnational evaluation and indicate strategies that might be used to overcome them.
- To contribute to greater awareness, transparency and compatibility within European higher education.

The TEEP project was initiated and supported by the European Commission through the SOCRATES programme with ENQA responsible for its delivery. A Planning Group chaired by ENQA included representation from the European University Association (EUA), the National Unions of Students in Europe (ESIB), and the Commission. The project was overseen by a Management Group, chaired by ENQA, and consisted of representatives from each of the participating quality assurance agencies, the

subject area Thematic Networks' chairs, and an EC representative. This group was involved in the design of the project, the selection of participating institutions and nominations for expert panels, and received and discussed intermediate reports. A Project Group, composed of staff members from the participating agencies, oversaw the details of the organisation and implementation of the project and acted as secretariat at the site visits. Expert panel members contributed to the relevant subject reports but the Project Group was responsible for their final preparation, after the institutions and experts had checked on the accuracy of relevant details and recommendations. The Project Group also prepared the draft methodological report. Representatives from the institutions, the subject experts and Management Group all contributed to the final version of the methodological report (for their contribution, see chapter 6).

1.2.1 Anticipated benefits from TEEP

The likely benefits from TEEP should include:

For European higher education:

- A method for transnational evaluation building on predefined criteria which are commonly agreed, which have been tested and which offer a dimension of transparency and comparability to the quality of programmes across borders;
- A contribution to the development of each subject on the basis of the recommendations of the experts and good practice from comparable programmes in other countries;

For the participating institutions:

- The opportunity for each of the participating institutions to promote both their institution and their programmes;
- The opportunity to receive feedback as a contribution towards improving their quality assurance culture;
- An opportunity to share experiences between programmes and peers, and the possibility of establishing networks to assure continuous improvement of programme quality.

¹ The term 'full programme' should be understood as covering both the first and the second cycle leading to a second cycle degree, in this case leading to Veterinary Science master degree.

1.2.2 Report structure

The methodological report is structured in six chapters. The first chapter introduces the TEEP project. This is followed by chapter two that consists of an executive summary of the overall results of testing the TEEP methodology.

Chapter three provides the reader with a short comparative view of the main findings of the three subject-specific reports concerning 1) The level of implementation of the first and second cycle degree structure, including correspondence with the Dublin descriptors; 2) The extent to which the programmes have formulated and used definitions of competences and learning outcomes, including knowledge and applicability of the TUNING crite-

ria; and 3) The level of implementation of quality assurance processes in the programmes.

In chapter four the method and its different elements are discussed in more depth. Thus, the reflections on the experiences of the different methodological elements of the transnational evaluation are based on feedback from the expert panels, the programmes and the Project Group's experiences from the project.

Finally, in chapter five applicability of common criteria in TEEP are discussed.

The report is completed with chapter six, in which the Management Group, the experts and representatives from the programmes reflect on their experiences with the TEEP project and methodology.

2 Executive summary

TEEP followed upon earlier initiatives in trans-European evaluation, such as the European Pilot Evaluations of 1994/95. TEEP was an ambitious as well as a realistic move forward in methodological experiences exactly because TEEP took into consideration these earlier cases, as well as the now-developed principles for quality assurance in higher education institutions. TEEP must also be seen in the parallel context of the ‘Quality in universities project’ organised by the European University Association; similarly an initiative supported by the European Commission.

The objectives of TEEP were to

- further develop a method for transnational external evaluation, building on experiences such as the TUNING Project and the Dublin descriptors developed through the Joint Quality Initiative, using common criteria as the basis of an evaluation process in three different disciplines
- identify potential obstacles to transnational evaluation and indicate strategies that might be used to overcome them
- contribute to greater awareness, transparency and compatibility within European higher education.

The overall conclusion of this report is that these objectives have been satisfactorily met.

Conclusions on testing the use of common criteria

One of the TEEP project aims was to test the use of common criteria. In general, the use of common criteria was important in order to ensure that all of the programmes were evaluated through a single approach. Across the three subject areas, however, the experts and programme representatives have met some difficulties in understanding and interpreting the criteria as these had been set out in the project manual. This was especially the case with regard to articulating competences and learning outcomes, and some aspects of quality assurance.

The conclusion on the applicability of the criteria is that it depends on their formulation and ‘readability’, and the extent to which they can be related to a nationally, as well as an internationally, accepted threshold. Further, their applicability also depends on the extent to which programmes have developed and implemented those aspects of quality assurance covered by the criteria. The criteria can have important roles in stimulating and supporting such developments where they are appropriate within national and institutional contexts, and their implementation relates to desirable or necessary aspects of quality assurance.

Conclusions on obstacles to transnational evaluation and measures to overcome them

One general conclusion is that the process would have benefited from an extended timeframe at certain stages of the process.

The project has further been challenged throughout by the budgetary constraints, particularly relating to travel arrangements. It would have been a benefit for the process if there had been more opportunities for the expert panels to meet and discuss, for instance during the drafting stage of the subject-specific reports.

Another general conclusion relates to methodological implications in conducting an evaluation in one language – as the language for most of the participants will be their second language. The project demonstrates the importance, and additional value of including within each panel at least one member who speaks the local language and ideally also has knowledge of the national and local contexts in which the evaluated programme operates. Further, it is important that the experts represent different profiles of experience and are drawn from different countries.

The project's contribution to greater awareness, transparency and compatibility within European higher education

The overall conclusion is that the participants in general found TEEP an interesting and stimulating project. The inclusion of institutions and experts from a variety of countries has helped to contribute in a small, but no doubt significant way, to a greater sharing of knowledge about approaches to quality assurance within European universities.

The publication of the TEEP methodological report and subject reports should therefore contribute to greater transparency and increasing awareness of the potential and importance of compatible approaches to quality assurance within a transnational framework. In this way TEEP has provided a valuable insight into aspects of the implementation of the Bologna and Prague processes, particularly at the level of academic programmes.

3 Results from the subject-specific reports

A comparison of the three subject areas at the European level has been developed through the subject reports, in accordance with the three themes identified within the TEEP manual – educational context, competences and quality assurance. The following chapter gives an overview of the main conclusions across the three subject areas concerning: 1) The level of implementation of the first and second cycle degree structure, including correspondence with the Dublin descriptors; 2) The extent to which the programmes have formulated and used definitions of competences and learning outcomes, including knowledge and applicability of the TUNING criteria; and 3) The level of implementation of quality assurance processes in the programmes. It is recommended that the three subject-specific reports be read in order to gain a comprehensive view of the evaluation of the 14 programmes – the evaluation of each programme should be read in their specific subject, national and institutional context. The three subject reports can be found on the ENQA website².

3.1 Degree structure and definition

One of the criteria evaluated through the transnational project was the extent to which the programmes have formulated and established a first cycle degree programme. The evaluation attempts to establish whether the programmes have formulated goals for the first cycle degree, and to what extent these formulations may match the ‘Dublin descriptors’ for the first cycle degree.

The extent to which the evaluated programmes have implemented a first cycle degree programme

varies considerably across the programmes. The Veterinary Science programme is a full programme leading to a degree that is recognised as equivalent to second cycle degrees. In contrast, the evaluated Physics and History programmes are almost all now in the process of implementing the first and second cycle degree structures, but they are at different stages of development.

The new History programmes within the TEEP evaluation are of 3 or 4 years in length and most are planned within the Bologna concept of first and second cycle degrees. Most of the Physics programmes have established first cycle degrees of 3 years’ duration; three of the programmes have just been re-structured according to the Bologna model. One of the Physics programmes does not offer the first cycle degree, but an integrated five-year programme leading to a second cycle degree instead.

It is evident that the extent to which the evaluated History and Physics programmes have implemented the first and second cycle degree structure is related to the commitment of the respective countries with regard to the Bologna process. It is not surprising that in those countries where the first and second cycle degree structure has become a part of the government’s higher education policy or regulation, universities are undertaking the implementation of this new structure.

Compared with the evaluations of History and Physics, the evaluation in Veterinary Science was carried out at second cycle degree level. This is a consequence of the present situation of a “long unique cycle” degree in this discipline. The two cycle structure, as it is defined in the Bologna Declaration, is not applied in these programmes. Moreover the evaluation showed that it would not be easy to formulate a shorter or intermediate degree without affecting the whole structure in the education of Veterinary Science.

² Available at www.enqa.net/pubs.lasso.

It should be noted, however, that the Veterinary Science programmes had characteristics of both first and second cycle degrees. The Veterinary Science programmes are clearly oriented towards the labour market, but at the same time the education of a veterinarian requires an amount of training that is evidently equivalent to a second cycle degree. All of these aspects open a relevant question about whether the new European framework is suitable for those professionally oriented degrees, such as Veterinarians or Medical Doctors, which require longer education programmes in comparison with other disciplines in which the degrees may be more or less targeted at a labour market.

Some of the evaluated first cycle degree programmes in Physics and History have formulated specific level descriptors for the first cycle degrees. The range and balance of competences appear to match to some extent those included within these descriptors. The competences formulated for the first cycle degrees mainly aim at giving an academic grounding in the subject leading to the second cycle and/or doctorate degree and prepare for the labour market. It is apparent that the relationships between the aims and structures of the first and second cycle degree programmes are being discussed, and that the TEEP project has contributed to this discussion.

The evaluation has shown that the extent to which the programmes have formulated specific aims for the first cycle degree varies considerably. The evaluation of the History programmes shows that most have identified explicit aims for first cycle degree programme and some are planning articulation with second-degree awards. Three out of the five institutions are delivering new programmes in line with Bologna but have yet to produce graduates.

In Physics one of the programmes has formulated explicit aims, stating that the first cycle programme leads both to employment and further study; the other programmes have not explicitly formulated their aims for the first cycle programme. For these degrees, it is implicit that the first cycle degree is the first step towards the second cycle or PhD degree. The extent to which the programmes have formulated specific aims for the first cycle degree seems, however, to depend on the interac-

tion with the labour market and whether labour market representatives have been involved in formulating the expected learning outcomes.

3.2 Competences and learning outcomes

The term ‘competence’ for the purpose of this project refers to a large extent to the outcomes of the TUNING project. The evaluation of the programmes looked to see how far the definitions of both subject-specific and generic competences have been used in the formulation of courses and programmes, and the extent to which these competence definitions have been used by staff to develop a shared understanding about the delivery and expectations of the programme, and also to what extent these have been communicated to students so that they know what is expected of them. The evaluation also looked at whether or not teaching and learning methods, as well as assessment methods were designed to support the development of the desired competences.

Although limited in its scope, this project has demonstrated that there are apparent differences in the extent of awareness about and knowledge of the TUNING subject-specific and generic competences amongst relevant academic and administrative staff, both across programmes and across disciplines. Furthermore, there are differences regarding the extent to which the programmes have formulated and used (or planning to use, or not) the TUNING descriptors.

Consideration of ‘competences’ has proved a useful focus for discussions within TEEP.

This has been the case on a generic level; in the discussion of the sort of abilities and approaches that are relevant to students reaching the end of the first cycle, but at the same time the particular aspects relevant both within and between disciplines are apparently different.

In History the subject-specific competences from the TUNING work are beginning to be recognised as relevant and being used by some of the staff within the departments visited. The TUNING criteria are seen as external reference points that can

assist in the development and articulation of programmes and their expectations. However, those who had considered them to be less relevant saw the TUNING criteria as generic competences; they are regarded as too numerous to be of real value. It was apparent that in the majority of cases the staff found it easier to relate to the subject-specific competences than to the generic ones.

The extent to which the Physics programmes are familiar and employ the competence terminology varies considerably. Two of the programmes have formulated both subject-specific and generic competences at programme and course level. For the remaining three programmes the competence terminology is unfamiliar and therefore not actively used.

In Veterinary Science, due to requirements of EU directives, national regulations and other European developments, the programmes are defined in order to deliver all subject related competences through compulsory subjects. At present, the programmes have defined a set of generic skills for the Veterinary education, but the balance of subject and generic competences is to be modified.

Another important dimension of the TEEP criteria and the application of competences is the extent to which teaching and learning strategies and assessment methods support the development of both subject-specific and generic competences. Although some of the History and Physics programmes have not yet explicitly formulated expected competences or communicated them effectively, the teaching and assessment methods do seem to support the development of both subject-specific and generic competences. There seems to be an implicit understanding amongst the staff as to what the expected competences are; however, these have not yet been made explicit to students in all cases. Also in Veterinary Science, the teaching and learning methods permit the achievement of subject related competences. However, in some areas a development of teaching methods would facilitate improved acquisition of skills and attributes.

Although Veterinary Science programmes have not participated in the TUNING project, they are accustomed to think in competence terms and to reflect on external reference points such as the

TUNING descriptors. The exercise and the vocabulary are familiar to the programmes since the European Association of Establishments for Veterinary Education have formulated similar transnational common core competences.

This was not the case in History and Physics where the evaluated programmes had some problems understanding and applying the terminology of competences and learning outcomes as set out. A substantial effort will therefore still be needed in order to ensure a shared vocabulary and culture of internal and external reference points, such as the TUNING outcomes and competences.

3.3 Quality assurance

The quality assurance criteria used within the TEEP project included consideration of whether programmes formulate their quality assurance procedures to ensure that the programmes remain current and valid in the light of developing knowledge within the particular discipline and its practical application, and the extent to which the aims and intended outcomes of the programmes remain appropriate to factors such as changes in student demand, student entry qualifications, employer expectations and employment opportunities.

From the documentation it appears that many of the evaluated programmes are in an early stage of formulating and implementing a general quality assurance strategy. The extent to which the programmes have a quality assurance culture seems to depend on national traditions of internal quality assurance mechanisms in higher education programmes. It should be noted that the professional dimension of Veterinary Science degrees, jointly with the existence of a European directive, have encouraged the implementation of some wider quality control measures.

The TEEP evaluations indicate that the quality assurance activities of the programmes mainly focus on course evaluation. A very limited number of the programmes have established a coherent framework for quality assurance that includes a broad range of quality assurance activities e.g. programme evaluation, course evaluation, alumni surveys etc.

The extent to which the programmes collect systematic feedback from stakeholders is low. In terms of curriculum development very few of the programmes have established, for instance, a systematic procedure for feedback from the labour market at programme level. Neither do the programmes collect feedback from the graduates on a systematic basis. Mostly, feedback is provided on a sporadic basis and depends on personal relations.

It has been evident that few of the programmes systematically gather and use information on student progress and graduate employment. According to some of the evaluated programmes, the exer-

cise of collecting statistical information on student progress for the transnational evaluation has been a very valuable experience, making the programmes more aware of their student population.

Though limited in its scope, the TEEP project appears to have made a significant contribution in stimulating discussions about and recognition of the need for the programmes to develop explicit quality assurance strategies. The programmes visited commented that they recognise that quality assurance strategies are necessary for the future and intend to build on their TEEP experiences.

4 Methodology

This chapter provides a description of and reflections on the methodological elements included in the evaluation. The evaluation model follows the European Council Recommendation from 24 September 1998 on European Cooperation in Quality Assurance in Higher Education where a so-called four-stage model for good practice on quality assurance is introduced. The four stages of this evaluation model include: 1) autonomy and independence, in terms of procedures and methods concerning quality evaluation, both from government and from institutions of higher education; 2) internal self-examination 3) externally-composed element based on appraisal and visit by external experts, and 4) the publication of a report.

The four-stage model is today generally accepted as the shared foundation of European quality assurance, and it also has a prominent place in the Council Recommendation of 1998, and in the criteria for ENQA membership.

The following chapter presents the process of the TEEP evaluation and the main methodological choices made. Thus, the Project Group reflects on the experiences of the different methodological elements of the transnational evaluation, which are in part based on the feedback from the expert panels, the programmes, and in part based on the Project Group's experiences from the project. Each section concludes with a summary of 'lessons learned'. The Project Group hopes that these reflections will be of value to future transnational evaluations.

4.1 Organisation

4.1.1 Process

The TEEP project started in June 2002 and was concluded with this methodological report. The whole process is illustrated in the figure shown below and described in detail in the following sections.

After setting up the organisational framework of the programme – consisting of the Project Planning Group, Management Group and Project Group, a **call for programmes** to participate within each subject area was made. The project required five institutions/programmes from different countries for each of the subject areas. A project plan and an introductory letter were sent to all the programmes participating in the CLIOHnet, EUPEN and EAEVE-network (the three Thematic Networks). A parallel process involved the **writing of a manual**, which contained guidelines for self-evaluation and formulation of the criteria that would be used by the panels in their evaluations.

The next step was the **selection of the participating programmes** by the Management Group. The manual was sent to the selected institutions in September 2002 and they were invited to participate in a **two-day working seminar** in Brussels at the beginning of October. Here the programme representatives had the opportunity to work together with the manual before starting their self-evaluation processes.

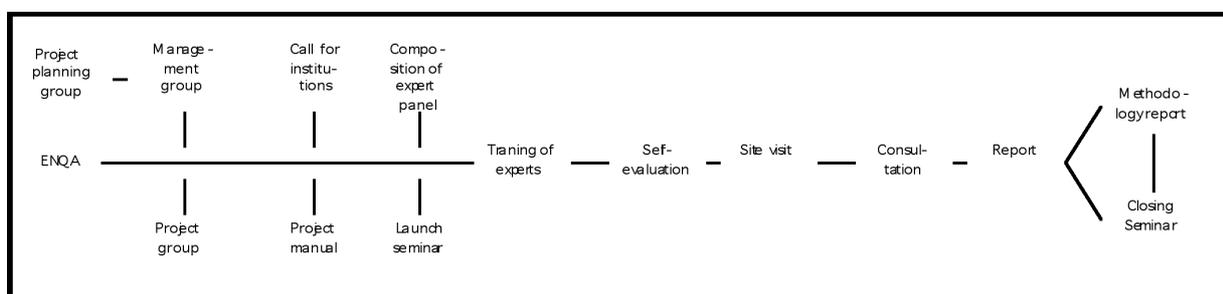


Figure 1. The process

From October to December each of the programme teams conducted their **self-evaluations**. During the self-evaluation period the **experts were identified**, approved by the Management Group and the programmes they would be evaluating, and then recruited. As many of the expert panel members as were available met in Brussels in January 2002 to undertake inductions and discussions about their roles and responsibilities. Following the submission of the self-evaluation reports (by December 2002 – January 2003) **site visits** by groups of 3–4 experts drawn from the relevant panels were organised for the period February to April 2003.

Following the completion of the site visits, the three subject reports were drafted in May 2003 by the three national agencies respectively, drawing on the extensive notes taken and on the panels' conclusions made during the site visits. These draft reports were sent to the experts for their consideration and comments.

In June 2003 the three draft subject reports that had been approved by the relevant experts were sent in **consultation** to the institutions for factual comments. **The three subject reports** were published in August 2003 (on the ENQA website³).

A methodological report coordinated by ENQA and prepared by the Project Group was drafted from June to August 2003. The main results of the project were presented to the ENQA members at the ENQA General Assembly in September 2003. A **closing seminar** was held in October 2003 with members of the expert subject panels and representatives from the evaluated programmes experts invited to contribute their final reflections and formulate their recommendations for future projects. The project completed with the publication of the **methodological report**.

One of the overall conclusions concerning the process is that the project would have benefited from more time, particularly at certain stages of the process. The period from institutions receiving the invitation to participate to the completion of the final report was only 8 months. Considering the project involved a transnational evaluation covering three

different subjects and 11 different countries and experts participating from an even higher number of European countries, this may be considered a quite extraordinary achievement. Many of the experts' and programme representatives' reflections refer to the time frame of the project. These comments will be presented in connection with the discussion of the different methodological elements in the following sections.

4.1.2 Scope of evaluation

The transnational evaluation covered programmes in the three subject areas; History, Physics and Veterinary Science. The European Commission was responsible for the identification of the three subject areas. The advantage of the three subject areas is that they represent both arts and science subjects, whilst Veterinary Science also represents a professionally oriented subject. It was considered more of a challenge to test common criteria across different subjects than if the programmes were from the same or similar areas.

From a methodological point of view it is interesting to compare more general degrees to professional oriented degrees. However, the evaluation has also been complicated by choice of evaluating Veterinary Science. The Veterinary Science area has a long-standing tradition of international evaluation and it was therefore difficult to recruit institutions to volunteer for the project. Those institutions that participated found it interesting and beneficial to compare the TEEP method to the EAEVE method.

In Physics and History the scope of the evaluation was from the outset confined to the programme level – more specifically to the first cycle degree or equivalent. However, the intention to focus on the first cycle degree was complicated by the fact that Veterinary Science, due to European Directives, is not committed to implement a first cycle degree. Therefore, it was decided that it would be more appropriate for the full Veterinary Science programme to be evaluated.

The focus of the evaluation on the level of the first cycle or equivalent was motivated by the emphasis in the Bologna Declaration concerning the application of a transparent system of qualifications in higher education based on two cycles, and the

³ Available at www.enqa.net/pubs.lasso.

increasing interest in the development of first cycle programmes.

Furthermore, the background of previous experiences with transnational evaluation and the TUNING project suggested that the programmes would be more likely to have greater common content during their first cycles than during the second cycle when there is a higher degree of specialisation. The greater proportion of common content was considered beneficial as the evaluation covered a range of countries as well as the different subject areas.

4.1.3 Recruitment and selection of programmes

As the project was based on cooperation between ENQA and the three Thematic Networks, invitations to participate within the project were generated through the members of the Thematic Networks via the lists on the network websites. It was an advantage for this pilot project that the recruited and selected institutions were part of a well-functioning network as the information flowed smoothly through existing channels of communication and the project benefited from existing links, understandings and mutual trust.

The disadvantage of only including network members was that the effects of the project were perhaps more limited. It may be assumed that within Europe the network institutions are amongst those that would already be more likely to be aware of the mechanisms for explicit quality assurance and their importance. It would probably be beneficial for a project of this type to reach out to a wider range of other institutions interested in European evaluation and exchange of information and practice. Furthermore, by inviting all European universities to participate in such a project the scope for selection would also be broader and it is likely that the effects of the project would spread to more programmes.

The Management Group using the following criteria selected the programmes:

- Involve as close to 15 different countries as possible with five different countries in each subject group;

- One country from Central and Eastern Europe should be included within each subject;
- Participation should be based on each institution's own initiative;
- There should be encouragement for inclusion of some institutions without former external evaluation experience;
- Participating programmes should have a minimum of 100 enrolled students;
- Institutions that had not participated in the TUNING project should be encouraged to participate.

An overall consideration of the geographical spread was also included, with a North – South-East dimension, and a small country – large country dimension considered within the selection process.

In order to ensure that the programmes evaluated had a sufficient number of students, programmes with a total enrolment (not per year or per cohort) of not less than 100 students were invited to participate. As the project builds upon the TUNING experience, it was interesting to invite both institutions with and without TUNING experience to participate.

The selection of institutions / countries was complicated by the fact that there were different levels of responses from the three subject areas. In Physics many participants volunteered whilst there were less in the other two subjects.

As mentioned in 4.1.2, the fact that the Veterinary Science area has a well-established international evaluation system was the main reason that very few Veterinary Science institutions volunteered to be included within TEEP, and in the end only four Veterinary Science programmes were evaluated. The scope for selection was therefore narrow and, as an exception, one of the selection criteria had to be adjusted and as a consequence two EU Associated countries participated in the Veterinary Science evaluation.

The geographical criteria were valuable in order to ensure a European perspective. In all three subject areas Northern, Southern and Eastern Europe were represented. An important dimension within the pilot project was in testing whether the same criteria could be applied across different European

countries despite their different educational traditions. The three reports together provide a case study with a perspective of 11 different European university systems and a provisional insight into where the different countries are in the Bologna process.

Lessons learned

- How to engage with a wider spectrum of institutions, and not solely institutions active in the Thematic Networks, should be considered.
 - The appropriateness of choosing a subject area that has its own quality assurance system should be considered.
 - The geographical criteria were important in ensuring the European dimension.
-

4.1.4 Budget

Throughout the process of the project the institutions and the experts have shown a significant commitment to participate in the evaluations. From the experts' point of view, their commitment has been challenged by the budgetary constraints of the project.

All the panel members participated without receiving a fee or other compensation for their time or work. All travel, board and lodging for visits and seminars were paid by the project but due to European Commission rules of economy class travel the members of the expert panels had inevitably to stay over a Saturday night on some occasions because of airline restrictions.

The fact that the site visits were carried out during the weekends and over a short period meant that the project to a large extent relied on the panel members' good will and enthusiasm. Whilst many were interested in this new type of project, it would be unreasonable to expect that a sustained and extensive international evaluation programme could be established and supported through such goodwill and generous free time of institutional staff and visiting experts. Additional time allocated for the overall TEEP project must be regarded as essential in the planning of any further transnational evaluations based on TEEP.

It is the view of the expert panel in Physics that the number of expert panel meetings was insufficient. Both the time and budget constrains limited

the possibilities for the panel members to meet and discuss sufficiently the preparation and recommendations.

Lessons learned

- Transnational evaluation should not normally be conducted on the timescale used by TEEP; additional time should be included
 - Time schedules and budgets should include a meeting for the panel members to meet and discuss the draft report and recommendations.
-

4.1.5 Language

It was decided that the transnational evaluation should be conducted in English. All self-evaluation reports, site visits in general, and panel discussions were in English. Practically it is necessary to use a common language in a transnational evaluation.

Methodologically there are some implications conducting an evaluation in one language – as the language for most of the participants will be their second language. It has been essential for the project that the experts and secretaries participating in the evaluation had good English language skills. The same is the case for the (majority of the) self-evaluation groups and the student and staff interviewed at the site visits. However, the extent to which the evaluated programmes use English as a means for communication varies.

In general, the staff and students at the site visits had a good command of English. However, this was not always the case. Fortunately, it generally proved possible through both design and coincidence to find shared languages other than English when necessary. The project did however demonstrate the importance and additional value of including within each panel at least one member who spoke the local language. The TEEP evaluations would not have been as complete had they been restricted only to those who could contribute in the shared language (English).

Another problem connected with the evaluation was the availability of documentation about the programmes in English. In some cases material describing a programme was only provided in the local language and in others the programme team / institution had contributed additional (scarce) resources

to have the material translated. The language issue thus set limits for how a transnational evaluation can study a subject in depth.

A last problem arising from conducting an evaluation in a foreign language is that lack of sufficient knowledge of the local language may impact on the extent of knowledge and understanding of the roles played by national, cultural and historical contexts. It is therefore an advantage to include within each panel an expert who has some knowledge and understanding of such specific contexts and their current impacts. In the Veterinary Science evaluations a local graduate participated in the initial meeting of the expert panel before the site visit. This proved of considerable benefit to the project as the expert group gained a good first hand impression of the programme context before the actual site visit. This meant that more time was available for in-depth discussions instead of gathering factual information in order to understand the programme context.

Lessons learned

- The primary language used for such a project should be considered carefully.
 - A common language is important for ensuring communication; practically English has been most useful.
 - It is valuable to have at least one expert who has some competency, if not fluency, in the local languages.
-

4.1.6 Selection and composition of the expert panels

4.1.6a Criteria for selection and composition of expert panels

When the expert groups were being selected and appointed it was emphasised that the expert groups should each represent different profiles and countries, thereby supplementing each other in different European aspects of quality assurance within each specific subject. It was decided that each expert panel should include at least one subject-area specialist, an expert with TUNING experience, an expert with managerial experience in the subject field, and an expert with didactic-pedagogical experience, and that different specialisations should be represented.

It has been an advantage for the project that some of the experts have had extensive experience with quality assurance reviews and / or of setting up quality assurance systems. Both institutions and experts have gained a mutual insight into different European quality assurance cultures through sharing such experiences.

In establishing the panels it was noted that the chairperson should have the necessary qualifications to fulfil the role and lead the panel. Thus, the chairperson should for example have a good subject reputation in the field being evaluated, have managerial experience, and most importantly have credible experience with quality assurance and preferably external reviews.

Other fundamental criteria applied to the appointment of members were:

- Members should speak and write the working language (English) fluently;
- Members should not be a resident of or have the same nationality as one of the countries participating in the particular evaluation;
- Members should not be directly connected to any of the institutions participating in the evaluation of the particular programme;
- Members should not be involved in the board of a Thematic Network.

There was an intention that a student member should join each of the visiting expert panels. The student should be from the country being visited in order to contribute with national context knowledge. In accordance with the criteria, the student should come from another institution in order to be independent from the institution(s) being evaluated. The role of the student should be to focus on areas and questions relevant for the student body. In the Project Planning Group, the National Unions of Students in Europe (ESIB) took the responsibility for identifying the students through their national organisations. However, ESIB was met with unexpected difficulties in this task with the result that unfortunately only one visit included a student member.

Opinions have been divided as to the effects and consequences on the quality of the evaluations of including a student member. The History panel was for example not convinced that its work would have

been enhanced by inclusion of a student member. Some members of the Physics panel however regarded it as a weakness that the project was unable to include student members, as this should have been part of the method testing.

In Veterinary Science another model was used. Some countries only have one Veterinary Science programme and the independence criteria for the students were therefore unlikely to be met. Therefore, graduates of the evaluated programme were identified to meet with the expert group and facilitate the experts' understanding of the national context of the programme. This approach proved to be a good contribution to the evaluation and generated good experience.

4.1.6b Expert panel

Due to the budget and time constraints of the project it was decided that the expert panels for each site visit would be drawn from a pool of 8 members. The visiting panel to each institution consisted of 4 members. For continuity, the panel chairperson and secretary participated in each visit.

The ideal number of experts for a site visit is open to discussion. The experience of the project shows that a panel of 3–4 members proved to be a good size, with the experts able to represent a range of different profiles and countries. In one case, due to late and unavoidable circumstances, one of the History panels included only 2 subject specialists and the secretary. This proved to be less satisfactory, and additionally placed a considerable burden on the team.

It was an advantage for the European dimension of the project that the experts represented different countries and educational cultures and were able to draw upon their different professional backgrounds when contributing to the project.

The appointment of a pool of experts from which each site visit panel could be selected was conceived as a solution to a practical problem. However, it might be a methodological advantage to have one expert group visiting all institutions; this would provide greater continuity, but might have reduced the range of experience and expertise contributing to the evaluation.

4.1.6c Division of labour between expert panels and secretaries

The quality assurance agencies have been responsible for the practical implementation of the evaluations and for applying relevant methods. Furthermore, they have acted as secretaries for the expert panels and have had the contact with the institutions and have drafted the report for the expert panels. The expert panels have been responsible for the academic content of the evaluations and for the recommendations and conclusions in the report.

This division of labour is well known from most of the quality assurance agencies in operation in the European Union countries allowing the experts to concentrate on the content and leaving the practical and methodological responsibility for the evaluation with the agencies, which have a professional expertise in evaluation.

The division of labour functioned well in History and Veterinary Science where as some of the members of the Physics team found that the lack of a common meeting between the experts to discuss the recommendations of report made ownership of the recommendations for all the experts difficult.

Lessons learned

- The experience of the project has shown that 3 or 4 members is a good size for the compositions of the expert panel for site visits, with the experts representing different profiles of experience and drawn from different countries.
 - It has to be clear to the experts what is the division of labour between the quality assurance agency and the expert panel when the experts are recruited.
-

4.2 Self-evaluation process

4.2.1 Self-evaluation manual

It has been valuable to have the same self-evaluation manual for a comparative project where a shared method is used for three different subject areas, and where the participants have experiences from different European national quality systems. The fact that the Management Group has been involved in the development of the method and the

institutions and the expert panels have participated in training, have provided a shared understanding of and commitment to the terms of the project.

The use of a single self-evaluation manual for the three subjects has generally ensured that the information provided by each of the institutions has been presented in a similar and consistent way, facilitating a comparison both across programmes within a subject and between subjects on a European level. Also, the provision of formatted tables for insertion of required quantitative data proved to be a valuable exercise. However, there were also some problems in applying national data into standard tables. It is therefore recommended that future transnational evaluations pay careful attention to the design and definitions of comparable indicators.

Furthermore, the responses from institutions and experts offered other suggestions for criteria and indicators, e.g. there was a single suggestion for addition of an indicator of the ratio of female and male students.

The institutions reported that they generally found the manual to be informative and useful and that it included all information and guidelines for the evaluation. However, the institutions also agree that the manual needs editing for consistency and clarification. The need for a revision of the manual and criteria is discussed in chapter 5.

Lessons learned

- A common self-evaluation manual is essential to ensure a shared understanding of and commitment to the terms of the project.
 - The manual used for the TEEP project would need revising prior to any further evaluations.
-

4.2.2 Launch meeting

A launch meeting was held between representatives of the self-evaluation groups and the project secretaries early in the process. Considering the pilot status of the project, such an early meeting was considered necessary to guide the institutions to provide comprehensive self-evaluation reports covering not only descriptions of existing practises but also reflections on these. Although there were variations in how analytical the content of the self-evalu-

ations reports were, the early initiation of the self-evaluation process is thought to have led to more analytical reports than would otherwise have been the case.

Throughout the process of the evaluation, the institutions and the Thematic Network chairmen have shown commitment to participate in the evaluation. It has been very important for the process that the Thematic Network chairmen and institutions were given the opportunity to comment upon the framework of the evaluation. The commitment from the institutions was also linked to the opportunity for the institutions to discuss and test the questions and criteria of the self-evaluation manual at the launch seminar. The launch meeting was also important in establishing a shared understanding and interpretation of the self-evaluation manual and self-evaluation process.

During the site visits, the self-evaluation groups expressed some uncertainties about interpretation of some of the questions and terms included in the self-evaluation guide. These had come up during later stages in the self-evaluation process – whilst it was possible to raise questions with the secretariat via, for example e-mail, it was noted that the representatives from the self evaluation groups would have found it more helpful to have a joint discussion meeting with the secretary later in the self-evaluation process. Whilst such a meeting might not have functioned well at an early stage in the process, a meeting during the self-evaluation exercise could have facilitated more discussion and clarification of questions, the terminology and not least the interpretation and application of the criteria used.

Lessons learned

- Consideration should be given to how to ensure a shared understanding and interpretation of the self-evaluation manual and self-evaluation process.
 - Regular contact with the institutions during the self-evaluation phase could help with interpretation of some of the questions and terms, and be valuable for the development of the self-evaluations reports.
-

4.2.3 Focused approach

The transnational evaluation was a ‘light’ version of typical national evaluations; it applied a focused approach structured around three pre-selected points:

- Educational context;
- Competences and learning outcomes;
- Quality assurance mechanisms.

The experience of using such a focused approach has generally been positive. It has provided the evaluation with an opportunity to make comparison between programmes more easily and more comprehensively. Further, the approach has been valuable in that it has provided focused self-evaluation reports that have been a stimulus to reflection within the institutions, and have been of help in structuring the site visits. It is perhaps worth noting that, with regard to the length and focus of the guidelines for self-evaluation and project, a broader evaluation with more areas would have been practically impossible both for institutions and experts to conduct in the timeframe given for both the preparation of the self-evaluation reports and the site visits.

A focused evaluation does, however, also imply that some issues are left out. The evaluation therefore presents only a part of the picture of the programmes compared with that provided by broader or more extensive evaluations. The Veterinary Science institutions and experts note that, although there is an EAEVE system to evaluate Veterinary Science faculties, the TEEP approach was appreciated as it introduced another method of evaluation. TEEP was lighter, less wide ranging, less rigorous in those areas that it addresses, and introduces some new aspects related to quality assurance.

The choice of the three areas for the focused approach was also an important aspect of the development of the TEEP method. The themes for self-evaluation generally seemed to be suitable although not all institutions could readily provide all of the requested statistical data.

Educational context as focus area

The focus on education context proved valuable; it would be difficult to conduct a transnational evaluation without a focus on both the national and the programme contexts. It was evident that the national legal systems have a significant impact on programme structure and composition, as well as on the organisation of quality assurance mechanisms. It would have been very difficult to understand the rationale underpinning the programmes without the information on national contexts. It cannot be emphasised enough that this focus is of vital importance when conducting transitional evaluation. Without it there is a significant risk that conclusions and recommendations will not be relevant to the evaluated programmes.

In the Physics evaluations the information provided about programme structure and content was not sufficient for the experts to fully understand the context in which the programme functioned. This influenced the quality of the evaluation. The Physics experts found it hard to assess the level and quality of the programme without information on the curriculum (or alternatively, detail of subject competences to be acquired). In their view, the manual should have been more explicit in asking the programmes to provide such material. However, the other panels did not express such a view, and indeed different arguments can be made on the extent to which the evaluation should go into detail on the programmes’ curricula and content.

In transnational evaluations a detailed consideration of curriculum and content could also have significant disadvantages. An in-depth examination of curricula would require individual institutions/programmes to provide details of their rationale for curricula design, strategies for teaching and learning, and how these are then reflected in approaches to (different types of) assessment, the setting and application of assessment criteria, and the rationale for progression routes and ‘hurdles’. All of these would need to be provided in the agreed language for the evaluation. All of this would be an immense work both for the participating institutions and also increase the workload of the panel substantially.

It could also be argued that it would be difficult for a transitional evaluation panel to assess programme quality through an in-depth study of the curriculum when the extent to which the curriculum can be varied depends to different extents on the consequences of national standards and requirements. The rationale of TEEP from the outset was that evaluations should be part of and support a process of development of quality assurance cultures within individual institutions through application of and reflection on the relevance of shared criteria.

Competences and learning outcome as focus area

The focus on competences and learning outcomes was an essential ingredient if the evaluation was to build on previous experiences with the TUNING project; this approach also provided an opportunity to consider the applicability of the recently developed generic descriptors for first- and second-cycle degrees (the so-called ‘Dublin descriptors’).

An aspiration associated with the creation of a European Higher Education Area is the establishment of comparable degrees with easily readable nomenclatures, within a system essentially based on two cycles. This requires the identification of relevant competences that are associated with each of the degrees. These are both subject-specific, but also require the identification of competences that are shared across programmes. The identification of such competences, and the development of shared understanding of their importance, should add to mutual trust and confidence about HE qualifications within Europe, and enhance those distinctive features of European cooperation that are closely linked with the transparency required to facilitate greater labour market mobility.

The relevance of competences and learning outcomes as a focus area was a subject of discussion among experts and institutions. One group considered that it was particularly valuable to focus on competence and learning outcome, as this provides the basis for the important change from comparing one syllabus against another to comparing the real outcome of the programmes. This then allows for different approaches and choices within a syllabus whilst still providing for the development of com-

parable competences.

Another group, however, found this approach difficult in concept and in operation. The majority of the institutions and experts were not familiar with the vocabulary of the TUNING project and it was therefore difficult to initiate a discussion of definitions at a sufficiently detailed level to take the comparative evaluation forward quickly. It proved difficult to convert what appeared to some to be an abstract terminology quickly into operational terms. For those institutions that were familiar with the competence terminology and had managed to use it operationally (i.e. provide explicit course and programme outcomes) it turned out to be an important tool in informing students of the content and goals of the programme.

Quality assurance as focus area

The last focus of the evaluation was on the programmes’ quality assurance mechanisms. In general, both institutions and experts agreed that this focus area was the one from which the programmes benefited the most. Several emphasised that the provision of statistics on student cohort and progress had been highly valuable to the programme.

It was also suggested that the focus on quality assurance should be seen as a way of helping institutions to set up a quality assurance system and make it effective. The fact that the evaluation is a transnational one, with a focus on quality assurance, can also be used as an instrument to help strengthen the arguments of individual universities in their requests for a greater focus on quality assurance within an international context, and in this way provide opportunity for greater mobility and international recognition.

Other focus areas

As mentioned earlier in this section, a focused evaluation does leave out some issues. Other focus areas such as research or information on infrastructure have been mentioned as important focus area by some experts and institutions and in case of Veterinary Science, practical training has been pointed out to be a focus area of relevance. However, a broader evaluation with more areas would have been

practically impossible both for institutions and experts to conduct in the time frame given for both the preparation of the self-evaluation reports and the site visits.

Lessons learned

- The focus evaluation approach makes comparison easier and the self-evaluation reports and the site visits more structured.
 - Focus on the national and local context is vital in a transnational evaluation.
 - It is important to have sufficient information on programme structure.
 - The competence area is difficult to approach if there is a lack of knowledge in the terminology
 - Quality assurance is an important focus area and helps the programmes in developing an effective quality assurance system.
-

4.2.4 Self-evaluation process

The preparation of the self-evaluation reports was designed to serve three distinct aims. It should provide:

- A framework to stimulate internal discussions of strengths and weaknesses related to the three themes (degree structure and definition; competences and learning outcomes; quality assurance) that are the foci for the evaluation. This was intended to assist the continuous improvement in the quality of the programme.
- Comparable documentation to be used by the panels of experts in their preparations, site visits, evaluations and reports.
- Comments on the utility of the criteria when applied to different programmes delivered within different national contexts.

The self-evaluation reports together with the information gathered during the site visits constituted the documentation for the evaluation.

In general, the programmes have pointed out that the self-evaluation process stimulated internal discussion about the quality of the programme. The evaluation was also seen as a starting point for (further) awareness of quality assurance and the self-evaluation process helped the programmes to discuss and see things in a new way. With more time

for the process the evaluations would have been more deeply rooted in the institutions.

However, some institutions also stated that they found the time too limited for self-reflection and for a broad participation by different stakeholders. The evaluation was nevertheless seen as a very valuable exercise that ought to be shared with the rest of the staff.

Lessons learned

- The self-evaluation process is important in stimulating internal discussion about the quality of the programme.
 - It is important that the time schedule allows time for reflection and for the participation of other stakeholders of the programme.
-

4.2.5 Self-evaluation group

A self-evaluation group was established for each of the programmes involved in TEEP. Under its chairperson the group prepared a self-evaluation report. The group consisted of five or more members and included representative from each of the relevant stakeholders at the programme level, including management, staff actively involved in teaching, students and administrative staff.

The self-evaluation groups generally functioned as anticipated within the project. Not only members of the programme's management team, but also students and staff represented in the self-evaluation group contributed to the site visits, and appeared to have an important ownership in the exercise and commitment to bring about the changes suggested by the overall process. Students and staff participating in the interviews at the site visits expressed an appreciation of the value of participating in the process.

However, a point for consideration in future evaluations is how to ensure that the results and ownership of the evaluation process could have a wide 'spreading effect' among the remaining staff and students who are involved in the programme. Often the self-evaluation group appeared to have prepared the self-evaluation report with little interaction with the rest of the programme. The manual should perhaps be more explicit in encouraging the

self-evaluation group to initiate a dialog with the rest of the programme in the self-evaluation process. If the information is not getting through to the rest of the programme there is a significant risk that the results of the evaluation will be of only formal value and not have any real impact or consequences.

Lessons learned

- A self-evaluation group with representatives of the different contributors and stakeholders in the programme should ensure broad ownership of the process and support widespread commitment to bring about any necessary changes
 - Consideration should be given to ensuring that the ownership and results of the evaluation process are widely spread among the staff and students who participate in the programme.
-

4.2.6 Quality of self-evaluation reports

The purpose of using a single self-evaluation manual was to facilitate a comparison across programmes from different national contexts, by seeking to ensure that the information provided by each of the institutions was presented in a similar and consistent way. The experience in this evaluation shows that even with such a framework there will be diversity among the self-evaluation reports.

Some self-evaluation reports were rather descriptive in nature; some thought this was due to the transnational context in which the evaluation was conducted. For the experts to be able to evaluate the programmes it is important that they are provided with a suitable framework. Since much of the self-evaluation document, and the discussions during the site visits, were concerned with the significance of a national and local context, the reports are not necessarily as evaluative as those produced from within a single country, where the context is recognised and understood. Significant consideration must be given within a transnational evaluation to the guidelines for the self-evaluation manual and the proforma tables for inclusion of comparable data sets. The different ways in which some of the institutions interpreted the instructions for the self-evaluation indicates that further development of the self-evaluation manual and report are required.

The descriptive nature of some self-evaluation reports may also have been influenced by the fact that a number of the questions raised in the self-evaluation requested descriptive answers. To counteract the descriptive tendency it would be valuable if questions encourage the self-evaluation panels to be reflective.

Lessons learned

- Transnational evaluations carry the risk of being descriptive due to the need for explanation of the national context
 - Consideration must be given to how the guidelines for the self-evaluation report and the proforma tables fit with national interpretations
 - The self-evaluation format should include questions that encourage reflectiveness.
-

4.3 Site visits

The self-evaluation was followed by site visits by the expert panels. The site visits took place in January–March 2003 and lasted 1,5 days per institution. All site visits were structured in a similar way, in accordance with a standard programme. The site visits provided the panel with an opportunity to invite the institutions to elaborate on unclear and less substantiated sections of the self-evaluation reports. At the same time, the site visits served to validate the information provided in the self-evaluation reports. Furthermore, the site visits allowed the experts to get a comprehensive and clear view of the programme through discussions and interviews with main stakeholders.

4.3.1 Organisation of the site visit

The length of the site visits was generally regarded as being appropriate. It would have been possible to complete each of the site visits in one day, but the inclusion of an extra half-day provided an opportunity to go into more depth on important issues. A one-day visit would have required that each of the interviews would have had to be shortened in order to allow for interviews with all of the stakeholders. The experience from the site visit showed that an interview of approximately 1,5 hour

is appropriate, and that less time would be insufficient to cover all of the necessary areas. Furthermore, it would not have provided the panels with the opportunity to prepare for each interview.

The organisation of the site visits, with separate interviews with different groups of stakeholders, served the intended purpose of validating the content of the self-evaluations reports by providing the panel with a clear picture of the opinions and perspectives of the different stakeholders.

The stakeholders interviewed at the site visits included: the self-evaluation group, the senior management, the teaching staff, students and graduates. The institutions were asked to provide 5-8 members for each group. This number proved to be appropriate and prevented potential difficulties arising where the experts outnumbered the group being interviewed. This was thought to be particularly important for the student group to ensure that they did not feel intimidated. A larger number than 8 would have been very time consuming and increase the difficulty of keeping a structured conversation.

Accordingly, it was emphasised before the site visits that an overlap of participants in the different sessions should not occur unless it was unavoidable due to the organisational structures relating to the programme. This proved to be difficult as there was often significant overlap between the self-evaluation group and the senior management team. In the Physics evaluation it was therefore decided to substitute the meeting with the management group with an additional meeting with the self-evaluation group. Both institutions and experts found this solution beneficial.

It had been anticipated that that the student representatives would be (randomly) selected according to the required programmes/specialisations. This also proved to be difficult in part because of the language barriers. The students were often, although not always, selected according to their ability to speak English. Institutions did a remarkable job in encouraging the students to participate and identifying students who met all the criteria and it was a general opinion among the experts that the interviews with the students were some of the most rewarding. In one of the evaluations the institutions had chosen to include both international students

on mobility programmes and a home student who had studied abroad within the group to be interviewed; this proved to be particularly valuable as they provided a comparative European perspective to the discussions.

The institutions were also asked to invite 5-8 graduates according to a number of criteria i.e. the graduates should represent different sources of employment. As in other areas of the evaluation the institutions showed a strong commitment to the project and provided the sufficient amount of graduates. It was interesting for the experts to gain an insight into the range of jobs the graduates were qualified to take up. Unfortunately, some of the graduates for the interview groups seemed to have been selected as a consequence of their outstanding curriculum vitae and they could therefore not be considered as being representative of the average graduates. The extent to which an evaluation with a focus on the first cycle degree benefits from interviews with graduates who are mostly doing their second cycle degree or PhD degree has been questioned. The evaluation might benefit more from interviewing both second cycle degree students and also graduates with a first cycle degree that are in employment.

One problem that should be considered is the extent to which the points of view of a small group can be considered representative. This is pertinent to the interviews with the students, teaching staff and graduates. Interviews with stakeholders at the site visit can only serve as an indication of this particular group of individual's reflections, not as being representative for the stakeholder group as a whole. In evaluations conducted by the national quality assurance agencies surveys on students, graduates and employers normally constitute a very important element of the evaluation of quality. Conducting surveys on transnational level would demand considerable resources. However, the surveying of more representative groups should be considered in any future transnational evaluation.

The Physics evaluation included a tour of facilities within the programme of the site visit. This was generally regarded as a positive experience in gaining an insight into the facilities. However, it was also noted that there is a particular and explicit need

to identify just what facilities the experts feel they need to see so that time aspects can be controlled effectively. Any tour must have a specific purpose related to supporting teaching and learning; it should not just be research equipment and computers that are not generally available to students.

Lessons learned

- 1,5 day for site visits with 1,5 hour for each interview has proved appropriate.
 - Separate interviews with different groups of stakeholders provides the panel with broad perspectives of the programme
 - 5–8 members for each interview group appears to work effectively
 - Overlap of participants in the different sessions should be avoided whenever possible
 - Interviewing international students on mobility programmes contributes to a comparative European perspective
 - A solution to the selection of students and graduates to provide a reasonably representative panel must be found.
-

4.3.2 Interview guides

For each site visit an interview guide was prepared; this provided the framework for the sessions to be held during the site visits. Reflecting the function of the site visits, the guide primarily contained questions related to those areas where more information or validation of information was needed in order to facilitate the subsequent comparative evaluation of the programmes.

The guide contained questions relevant to each of the groups to be interviewed, including some initial questions related to the self-evaluation process about how the various stakeholders had been involved in or informed about the process and the self-evaluation report. This was followed by a number of questions related to the content of the self-evaluation report and the focus areas.

In general the interview guides proved useful for the site visits. As the time for each interview was short the interview guide served to structure the interviews, and hence to generate the information needed to supplement the self-evaluation reports. The interview guides also functioned as a useful

instrument for the expert group and secretary jointly to prioritise the questions before the meetings.

A disadvantage of using relatively strict interview guides was that it countered the flow of the interviews to some extent and often meant that there was insufficient time to follow up on answers. Experience also showed that interview guides with too many questions limited the room for reflective aspects. However, the interview guides did ensure that all focus areas were treated equally at each site visit, thus facilitating the comparative element of the evaluation. Comparable, institution-specific interview guides are, therefore, recommended, provided that the length of the individual interviews is carefully taken into account when deciding upon the number and type of questions to be posed.

Lessons learned

- An interview guide is useful to assist structured interviews and to get the information needed to supplement the self-evaluation reports
 - In comparative evaluations the interview guides ensure that all focus areas are treated equally at each site visit, thus facilitating the comparative element of the evaluation
 - The length of the individual interviews should carefully taken into account when deciding upon the number and type of questions to be posed.
-

4.4 Report

TEEP resulted in publication of three subject-specific reports – one for each discipline. Each report was structured in two parts. The first part made a comparative view across the programmes according to the three focus areas. The second part contained programme reviews focusing on strengths and weakness of each of the programmes according to the criteria of the manual. This part also included recommendations for further development of the programme.

The character of the reports was more descriptive than initially expected. However, due to the different national cultures and systems, it was considered necessary to use a considerable part of the report to describe the national context of each of

programmes evaluated. This dimension of the project was underestimated in the project planning. As previously mentioned in chapter 4.2.3 on the focus area on educational context, this part is worth giving a higher priority in the planning of future transnational evaluations.

The first step in the process of report writing was that for each evaluation the secretary wrote an initial draft of the report. The experts received this

draft version with a request for comments. Unfortunately, the tight time schedule of TEEP did not allow an additional meeting of the experts, but the secretaries collected and reviewed comments and incorporated these into the final report. Subsequently, the institutions had the opportunity to comment on factual elements. Lastly, the report was sent to the experts for a consultation before publication.

5 Results and conclusion on the use of criteria

5.1 Criteria-based evaluation

In national evaluations of educational programmes quality is often assessed in terms of the extent to which the individual programmes achieve their own goals, and the legal regulations under which they operate. This approach is commonly referred to as assessing the ‘fitness for purpose’.

However, the goals of the programmes participating in this transnational evaluation, and the legal framework under which they operate, differ and therefore it is necessary to have a common framework of reference for the evaluation in the form of criteria. Generally, the common criteria have been useful in making such comparison possible. They have provided the three external expert panels and the programmes with a transparent basis for the evaluation of programmes across three different disciplines and the 11 countries.

At the same time, the evaluation has demonstrated that the formulation and application of the criteria must be carefully considered. The section below describes how the criteria have been formulated. This is followed by a critical discussion of the application of the criteria. Finally the Project Group concludes on the applicability of common criteria in TEEP.

5.1.1 Formulation of the criteria

The two sets of criteria relate to the two focus areas: ‘competences and learning outcome’ and ‘quality assurance’. The third focus area ‘educational context’ deals with the programme context and has served as a means to provide factual information about the programme; therefore no criteria were formulated for this area. The criteria have been formulated with reference to a number of different sources. Overall the objectives of the Bologna Declaration and the agreements reached at the 2001 Prague meeting have constituted one important ref-

erence point for the formulation of the specific criteria. Another important source has been the TUNING Project.

Under the focus area ‘competences and learning outcome’ the criteria are particularly concerned with the content of the programmes in terms of subject-related and generic competences, a terminology that was applied in the TUNING project.

Other criteria for ‘competences and learning outcome’ focus on the extent to which the first-cycle degree programmes (in the case of Physics and History), and second-cycle degree (in the case of Veterinary Science) corresponds with the formulated objectives in the Bologna Declaration. The criteria have been prepared on the basis of the Dublin descriptors for first and second-cycle degrees developed by the Joint Quality Initiative (<http://www.jointquality.org>). This developmental activity has been undertaken in line with the Bologna Declaration that proposes the introduction, within a European higher education space, of a system of qualifications in higher (tertiary) education that is based on two cycles. Each descriptor indicates an overarching summary of the outcomes of a whole programme of study. The descriptor is concerned with the totality of the study, and a student’s abilities and attributes that have resulted in the award of the qualification. In this way the descriptors are also connected to the set of the criteria that concentrate on the learning outcomes of the programme.

The set of criteria associated with the focus area ‘quality assurance’ was formulated to provide a basis for an analysis of the comparability of the systems and procedures applied by and on the participating programmes. They consider strategies, procedures and systems for quality assurance. One intention of the TEEP project was that that the criteria on quality assurance should assist the programmes in the development of quality assurance processes. The selection of these criteria has rested

upon the experience and knowledge that the European Network for Quality Assurance in Higher Education has gained from the implementation of numerous evaluations of higher education programmes.

The experts and programmes were asked to assess the formulation and applicability of the criteria used in the evaluation. In the self-evaluation, during the site-visits and in the aftermath, the programmes and experts have been very thorough in their reflections and feedback on the criteria. The following section, reflecting upon the application of the different sets of criteria, is based on the feedback from the programmes and experts, and the Project Group's own experiences with the project.

5.1.2 Applicability of the criteria

In general, the criteria have been important in order to ensure that all of the programmes were evaluated through a single approach, thus ensuring comparability in the assessment of the three subjects and 14 programmes. It is important to emphasise that the criteria were not being used for making formal judgements but were primarily used for testing their effectiveness in ensuring comparability. The criteria were directly linked to the focus areas and their related questions asked in the project manual. However, based on the feedback from programmes it is clear that this link could have been made more explicit.

5.1.2a Criteria for competences and learning outcome

In this set of criteria the programmes were asked to identify the establishment of programme goals, and reflect on the extent to which the goals were comparable to those encapsulated within the Dublin descriptors.

The evaluation revealed that there is considerable variation in the degree to which the programmes have formulated specific aims for the first cycle degree, and the extent to which they align with the Dublin descriptors. Some programmes have an explicit aim, stating that the first cycle programme leads both to employment and further study. Other programmes have not explicitly formulated their aims for the first cycle programme.

In general, the programmes and experts found it relevant to ask about programme goals. However, according to some experts and programmes the questions concerning goals for the programme were not formulated clearly enough. The manual operated with the terms 'goals', 'aims' and 'goals for competences' and the difference between these terms was not made clear to the programmes. It is suggested that the terms are defined more precisely.

The programmes were also asked to describe and reflect on the formulation of subject-related and generic competences, their consistency with programme content, and how the students may achieve the competences through the programme.

The focus on 'competences and learning outcome' in the evaluation was an interesting and a new experience for many of the programmes and experts. It was interesting to consider the generic competences and the subject-specific competences in regard to the discussion of which abilities and approaches that are relevant to students reaching the end of the first cycle. However, the application and understanding of the terminology were very different across the disciplines.

In History the subject-specific competences from TUNING were being used and recognised as relevant by some of the programmes. The TUNING criteria were seen as external reference points towards which the programmes can place themselves. However, the TUNING generic competences functioned less effectively, because they were regarded as too numerous to be of real value. It was apparent that in the majority of cases the staff members found it easier to relate to the subject-specific competences than to the generic ones. The competence terminology was further complicated by the fact that the TEEP manual did not adequately distinguish between competences and learning outcomes.

Across the three subject areas the experts and programmes had some difficulties understanding and interpreting this set of criteria. For many of the programmes, the use of the competence terminology was unfamiliar and therefore not applicable. Evidence from the project suggests that the dissemination and impact of the TUNING terminology in general, is limited and not necessarily widespread even within departments that include members of

staff involved in TUNING. The application of the terminology therefore seemed artificial and too complicated to some of those involved in the project and site visits. It can be concluded that a considerable amount of effort is still needed in order to ensure a wider recognition of internal and external reference points, such as TUNING and the competence-terminology, and acknowledgement of their value in articulating standards and ensuring their quality assurance.

5.1.2b Criteria for quality assurance

The experts and programmes in general found the criteria relating to quality assurance to be relevant. However, the extent to which they were being applied varied according to subject area.

The History programmes approached the quality assurance criteria with varying degrees of enthusiasm, commitment and support, and with different expectations about their continuing applicability. Some considered them as a starting point to be developed within the local context (after some initial scepticism). Others saw them as potentially authoritative within European developments and indicated that they would be seeking to continue their use in support of quality assurance changes, but that some detailed redrafting would be necessary. There was however a general recognition of the value of some/many of the criteria and, in particular, those related to the quality assurance role of self-evaluation.

In Physics the programmes generally appreciated the relevance of the focus areas on quality assurance. However, the programmes also noted some difficulties in applying all the criteria to their practise.

The professional dimension of Veterinary Science degrees jointly with the existence of a European directive has traditionally encouraged the implementation of some quality control measures, and the approach of the criteria appears to have been more immediately recognised and valued. The reflections obtained from the self-evaluation reports and expressed during the site visits noted that the programmes found that the TEEP evaluation introduced some new aspects related to quality assurance. The Veterinary Science programmes, however,

also noted some difficulties in understanding the definitions in the manual.

The Project Group suggests that the applicability of the quality assurance criteria is related to the maturity of the quality assurance systems of the programmes. Those programmes that have had experience in setting up systems that cover different aspects of quality assurance have found it worthwhile to be evaluated according to all the criteria. These have provided an 'added value' through comparisons to other European systems. However, for those programmes that are still in the process of building up a broad quality assurance system, only some of the criteria were applicable. For these programmes the remaining criteria might serve more as an inspiration for the continuing development of quality assurance processes.

Project participants have provided suggestions for re-phrasing of the criteria. Whilst these are not presented here in detail, it is anticipated that they will provide valuable help in future revisions of the criteria.

5.1.3 Conclusions

One of the aims of TEEP was to test the use of common criteria. The Project Group concludes that, at a general level, the applicability of the criteria depends on their formulation and 'readability', and the extent to which they represent a nationally as well as internationally accepted threshold. Further, their applicability also clearly depends on the extent to which the programmes have developed and implemented aspects covered by the criteria. The criteria have important roles in stimulating and supporting such developments and implementation where they are recognised and acknowledged as relating to necessary aspects of quality assurance.

The TEEP project has illustrated that there are considerable differences in terms of the educational cultures, national traditions and regulatory systems within which the individual programmes operate. Though limited in its scope, TEEP has therefore clearly demonstrated that with European higher education national systems and educational priorities still predominate. For that reason it is important that the differences in educational cultures are

carefully considered and the criteria are formulated in order to be flexible enough to allow for the impact of local and national contexts, legislation, and developments.

The project has also shown, however, that where national states have committed themselves to political objectives (aligned to the Bologna process) it is easier to reach a common interpretation. The project has in that respect provided a valuable insight into the condition for the implementation of Bologna and Prague at programme level.

It is important to emphasise that the criteria approach has provided the basis for making comparison possible. The common criteria have functioned as shared reference points, and ensured that the same topics were evaluated across the three disciplines and the 14 programmes. Any future transnational

evaluation project should seek criteria/reference points that are compatible with national and local contexts, and use terminology that is, to a large extent, familiar and useful for the programmes being evaluated.

Overall, it can be concluded that the participants in general found TEEP an interesting and stimulating project. Hence, many of the programmes have noted that they considered the evaluation very useful in providing the stimulus for extensive discussions about quality assurance in general and the nature and relevance of the criteria in particular. The TEEP process has thus proved to be valuable from a number of perspectives, but it has also provided an indication of further developments and improvements that would better support a criteria-based, transnational evaluation.

6 Reflections on methodology

TEEP is a pilot project with the objective of investigating the operational implications of a European transnational quality evaluation using common criteria. It has therefore been an essential part of the process to engage the institutions and the expert in the evaluation of the method used. The first five chapters of this report was drafted by the Project Group on the basis on comments from the institutions and experts and on the Project groups own experiences. The draft report was presented to the involved experts and institutions at a joint closing seminar in Brussel October 2003. The experts and institutions were invited to discuss the method of TEEP at the seminar. Furthermore they were invited to send their reflections in a written format to be included in the report. In the following section the comments are presented in a non-edited format. The Project Group has received comments from the Management group, the Physics programmes and the Physics expert panel.

6.1 Reflections from the management group

The Management Group of TEEP discussed the TEEP methodological report at its meeting 10 September. The overall judgement of the Management Group is that TEEP has succeeded well in reaching its objectives of testing a methodology for transnational evaluation in higher education and in discovering what obstacles are encountered and what improvements can be made – naturally within the framework provided by budget and other constraints. The Management Group considers the report to reflect fairly and precisely the strong and weaker elements in the TEEP process, but the Management Group has at the same time a number of recommendations for additions and revisions.

First of all, the Management group wishes to emphasize that all the phases of TEEP took place

in a dynamic context. Most institutions are now involved or have recently been involved in reforming their degree structure and courses study in accordance with the Bologna process; the TUNING I project was just closing as TEEP was launched; the final report of TUNING I became available several months later. As a consequence, the TUNING criteria had just been elaborated and were being disseminated in the institutions but knowledge of them was still limited. The Management Group reminds the reader that the Methodological report should be read keeping this dynamic context in mind.

Furthermore, the Management Group emphasizes that TEEP should be considered in the context of several other transnational projects. In this way the positive added value of TEEP can be brought into sharper focus. TEEP after all is in some aspects more transnational than earlier projects because transnationality is the basis for the identification of institutions, programmes and experts as well as quality assurance agencies.

In general terms, the Management Group accepts the need for a project to achieve the optimal balance between existing resources and overall objectives. Accordingly the Management Group is fully aware that the relatively limited resources available for TEEP have made it necessary to focus the project tightly, using a so-called “lighter touch”. This “lighter touch” has made the project possible, but has at the same time put restrictions on the applied methodology and on the scope and level of the evaluations. An in-depth review of the quality of programmes was not feasible and the scope could not for instance include an evaluation of the relation between teaching and research. However, because the “lighter touch” has been essential to the successful conclusion of TEEP and has provided credible evaluation results as well as being appreciated by the institutions, the Management Group suggests that the “lighter touch” be reflected upon in upcoming transnational evaluation projects in

order to investigate further the method's usefulness for evaluation and in developing a positive "quality culture".

The Management Group finds more problematic the extent to which the Dublin Descriptors or the TUNING criteria may be considered adequately tested. However, the Management Group points to the results of TEEP which is that the criteria have in general terms been important in order to ensure that all of the programmes were evaluated through a single approach.

At the same time it must be acknowledged that across the three subject areas the experts and programmes had some difficulties understanding and interpreting the set of criteria concerning competences and learning outcome and quality assurance. The Management Group agrees that the applicability of the criteria depends on their formulation and 'readability', and the extent to which they represent a nationally as well as internationally accepted threshold. Further, their applicability also clearly depends on the extent to which the programmes have developed and implemented aspects covered by the criteria. The criteria have important roles in stimulating and supporting such developments and implementation where they are recognised and acknowledged as relating to necessary aspects of quality assurance.

In any case the Management Group strongly suggests that it will be useful to provide feedback on experiences obtained from TEEP into TUNING II, especially in terms of analysing the methodology used and its feasibility. Further transnational projects will benefit from the work in TUNING II while at the same time TUNING II will benefit from the TEEP experience, continuing the positive dynamism evident during the TEEP project itself.

The Management Group considers the TEEP evaluation manual to be generally well prepared. However, the Management Group feels also that the definitions and terminology applied could have benefited from clearer expression beforehand. If that had been the case the participating institutions and programmes had stood a better chance of understanding the criteria in the same manner. In future projects the evaluation manual must provide exact

definitions of terminology and be drafted on the premise that most persons involved (experts, agency personnel and secretaries, staff and students of the self-evaluating institutions) will not be native speakers of English. For this reason special attention should be given to the simplicity and clarity of syntax and to clear definitions of terminology.

The Management Group takes note of the fact that Central and Eastern Europe institutions had a comparative disadvantage with respect to Western European institutions in terms of prior experience in transnational evaluations. For example, CEE institutions did not participate in TUNING although some were informed about it through the respective Thematic Networks. The Management Group considers this unfortunate but notes with approval the fact that they are more directly involved in Tuning II. CEE institutions must not only be included in further transnational evaluation projects, but they must also be expected to share fully in the methodological developments and projects, such as TUNING.

Concerning the selection of external experts transnational projects require that the appointed experts have knowledge of the different local higher education systems involved and the experts should also be selected according to criteria that include as wide a geographical distribution as possible. The Management Group makes a special recommendation for future projects by emphasising that the transnational aspect of the project must be very visible not only in the methodology but also in the composition of expert panels. This will be vital for the credibility and the acceptability of transnational evaluation. This is an issue not only of image but also of substance, as knowledge and experience of different national approaches will be necessary for all parties – including quality assurance agencies – involved in developing the European Higher Educational Area.

The Management Group would in principle have wished for TEEP to contribute even more clearly than it has to transparency and compatibility of quality of European higher education. However, as stated above, the Management Group accepts that the primary TEEP objective has been to act as a

methodological pilot to explore how the proposed methodology actually works in a variety of concrete situations. If the scope and level of the subject evaluation have necessarily been limited, this fact in no way detracts from the success of TEEP in reaching its primary objectives. In terms of methodology TEEP may have been slightly too optimistic in its initial ambitions. Still, the Management Group considers TEEP useful exactly because the results will help the planning of other projects.

To conclude, the Management Group recognises the importance of the fact that the final edition of the methodological report will include a further section reflecting the discussions to be held on 13 October 2003 when the institutions and the experts will meet in Brussels to share their views on the project.

6.2 Reflections from the institutions

6.2.1 Report from the evaluated Physics programmes⁴

We provide our answers to the four questions addressed in the guidelines of the Seminar, with a preliminary remark: the scope of the evaluation were the 3-year programs, but only one university had a real 3-year program in full operation, as others were in the process of completing their new 3-year programs and another had a 5-year program.

1. One of the TEEP objectives was the development of a transnational evaluation manual with common criteria applicable to three different subject areas. To what extent has TEEP fulfilled this objective? Identify possible weaknesses as well.

This objective was only fulfilled to some extent; this was due to several reasons including the following one. Indeed, the manual was rather unclear: e.g. the criteria given in the appendix were too numerous, sometimes redundant or obvious, other times confusing or not clearly enough formulated or too difficult to assess. In addition very little information was asked about the contents of the teach-

ing programmes and about the level of knowledge transferred to the students. As the manual was mostly devoted to the organizational and logistical aspects of the study programmes, any issue concerned with the link of the studies to research and to research-oriented training was missing. Of course, the part related to the topic of quality assurance was much appreciated.

2. The scope of the evaluation considered according to the three main dimensions: Educational context, competences and learning outcomes and quality assurance mechanisms. To what extent were the focus areas beneficial? Take into consideration the international dimension of the evaluation, and the third TEEP objective to contribute to greater awareness, transparency and compatibility, within European HE. What are your proposals to improve the evaluation approach?

The three focus areas were in general beneficial, as we found it useful to reflect and answer the related questions. However, we feel that the 'Tuning' criteria should not have been taken as a reference standard, because in the 'Tuning' pilot they were clearly intended to being considered only as a starting point. We estimate that the contribution of the TEEP project to awareness, transparency and compatibility within the European HE is quite limited. Surely the self-evaluation teams are now much more aware than before the evaluation, but little awareness has been created among the rest of the academic community, because we think that not many of our colleagues will read the reports (see also answer to question 4).

3. The TEEP process consisted of an internal and external evaluation. Could you underline good practices and improvement areas in both steps? (internal participation, site-visit schedule, calendar, meeting with stakeholders, the functioning of internal and external teams).

As far as we are concerned, the schedule of the site-visits was good, although the overall schedule of

⁴ Giovanni V. Pallottino, Dept. of Physics, University La Sapienza, Rome, Italy, October 20, 2003.

the project was too tight, particularly the phase devoted to finalizing the subject-specific report (leading to some factual error as indicated by the Vienna internal team). We also feel that the experts should have been more involved in the process, whereby the coverage concerning the teaching contents would have been improved. The experts should have been given the opportunity to visit all or most of the institutions in one group. Immediately after each site-visit time should have been foreseen for discussion and for writing up of their first draft report. In addition, in the group of visiting experts there should have been (at least) one expert who would be fluent in the local language and terminology, and who would have some knowledge of the local academic culture.

4. Reporting was done in two blocks: subject-specific reports containing the results of the external evaluations of the individual institution and a methodological report. To what extent are the reports useful tools for institutions? For the subject area at the European level? Identify strong and weak points of the reports and reporting procedure.

The reports are certainly useful for the visited institutions. But in order to be really effective, the reports need to be written in such a way that they would be worth reading. The crucial question is here: who in the academic community of our five visited institutions and in other European universities, will really read them in the form they are published now on the ENQA website only? As a matter of fact, both reports appear to be written by specialists in the evaluation discipline and to be only addressed to other specialists in this same discipline using the appropriate jargon, rather than to their natural intended target, i.e. the academic communities of the subject areas concerned.

6.3 Reflections from the expert panels

6.3.1 Comments on the TEEP Methodology from the Physics Expert Group⁵

The Physics Expert Group found the TEEP project interesting and stimulating. We were all pleased to have had the opportunity to be involved with the project and have benefited from working together and from the opportunity to learn about the institutions we visited. The project, and our participation in it, will probably help to the profile of Quality Assurance in our own institutions and in those we visited. However, we were disappointed that the contributions we could have made as experts were not fully exploited, and we would have been happy to have had more asked of us.

We were asked to look at four questions concerning the evaluation manual, the scope of the evaluation, the internal and external aspects of the evaluation, and the reports. We comment on these in turn below.

1. **The Manual.** We think that the manual as supplied was a reasonable first attempt at writing such a document and it shows that the approach of developing a manual with common criteria for a trans-national evaluation could work. However, the language was in many places ambiguous and unclear and much of the terminology would not be accessible to teaching staff. If the TUNING criteria are to be referred to, then it must be ensured that this terminology is familiar to the reader. A comprehensive glossary of all the terms used is essential – an attempt was made at this in the manual but this was not very useful and the definitions were not clearly written.

In order for such a manual to work well, it is crucial that there is no ambiguity in the language, so that the information supplied for different programmes is mutually consistent. In particular, the

⁵ Compiled by Prof R C Thompson, Imperial College London.

tables of numerical data in this manual were imprecisely defined so that even in the final report for physics there were still inconsistencies of interpretation between the five evaluated programmes. It was hard to compare programmes without the information being given in a consistent format.

2. **The Scope of the Evaluation.** We understand that for practical reasons it was necessary to limit the scope of the evaluation. The inclusion of Quality Assurance (QA) issues has helped to raise the profile of quality assurance in the expert group and in the institutions running the evaluated programmes, although this may be limited to the staff directly involved in the TEEP project and may not have spread effectively to all teaching staff. In general, the interpretation of QA in the manual is focussed on particular procedures and we felt that more emphasis on the outcomes of QA practices rather than the procedures themselves would have been fairer to the institutions and more helpful in general.

Our main concern here is that the interpretation of “Educational Context” should have been rather broader in the project. We would have liked to see more information provided on the Subject-specific Competences to be developed, which was requested in the manual but was mostly supplied only in a very general way. This would not be for direct evaluation of the details of the subject content of the degree programmes, but for the purpose of background information necessary for appreciating what the programme sets out to achieve. We would also have appreciated receiving some information on the research profile of the department, the provision of learning resources, the experience and qualifications of the teaching staff, and the involvement of the department in schemes such as ERASMUS (which is highly relevant to European mobility of students). This information would allow the experts to set the programme in context, especially as the student learning experience will be strongly influenced by the research environment.

3. **The evaluation process.** We found that the process of self-evaluation had a very positive effect on the institutions in general. It stimulated internal discussions about the programmes and a review of the teaching activities. However, this may have been limited to the members of the self-evaluation team in some cases, and these members may not have been representative of the whole department. The documents produced by the self-evaluation process are helpful and informative. For the external reviewers, they provide a very interesting insight into the evaluated programmes.

We were concerned about several aspects of the panel visit. In particular, we would have preferred to have a greater involvement in the evaluation process – both helping to set the agenda for the discussions and putting together conclusions and recommendations. The self-evaluation reports of all institutions should have been made available before the first visit took place. We felt that the visits should have been longer in order to allow for more internal discussion among the visiting panel. An initial meeting at the start of each visit to discuss the main issues to be addressed during the visit and to review the interview guide would have helped to focus the meetings with the institution. At this stage (or even before the visit), each member could have been assigned to a particular aspect of the evaluation so that he or she could concentrate on drawing together all the information on that topic.

We strongly recommend that at the end of the visit a further meeting of the panel take place, at which the preliminary conclusions and recommendations of the panel are discussed and formulated. We regard the lack of such a meeting as a serious weakness of the TEEP project. We suggest that at this meeting each member of the panel should present his or her conclusions on their assigned aspect of the review to the rest of the panel for discussion. An initial draft of the report for that visit should be prepared at this stage, and discussed with the self-evaluation group, before departure.

4. **The final reports.** We believe that the attempt to find a common language for the description of degree programmes at different institutions in different countries is a good one and that it has been generally successful. The physics report gives an interesting overview of the way that the programmes run in different institutions.

However, we are disappointed that the process of preparation of the report on the physics evaluation was not transparent. As a result of the way in which the report was prepared, it has resulted in a rather bland and repetitive report which does not contain as much substance as could have been possible. We believe that had the experts been more involved in the preparation of the report from the beginning it could have resulted in a more substantial report which the experts would have been able to identify with and which may have been of more use to the institutions. As it stands we do not feel that we as the experts have true ownership of the report. In addition, we doubt that it will be seen to be of much use to the institutions involved in the evaluation. A summary version, intended for wider distribution, would have had more impact.

In conclusion, we are pleased to have been involved in the TEEP project and we have all benefited from our involvement in various ways. It was unfortunate that the project was run at a time when some of the institutions were in the middle of a transition to a new Bologna-compliant programme structure, but this was unavoidable. We believe that the project will result in significant sharing of good practice between the evaluated institutions and our own institutions. TEEP was a pilot project. We see some good points in the project but we do not believe that it can be used for further trans-national evaluations without major revision. We have suggested ways of addressing some of the weaknesses we found. We strongly recommend that these suggestions be adopted in any future project along the lines of TEEP.

7 Appendices

Appendix A

History

History programmes:

- University of Bologna, Italy
- University Pierre-Mendez-France, Grenoble, France
- University of Coimbra, Portugal
- University of Aberdeen, Scotland (UK)
- University of Latvia, Latvia

Expert panel:

- Dr Colin Brooks, School of English & American Studies, University of Sussex
- Professor Steven Ellis, Department of History, National University of Ireland
- Dr Raphaela Averkorn, Universität Hannover
- Dr Taina Syrjämaa, Department of History, University of Turku
- Professor Tity de Vries, History Department, University of Groningen
- Professor Juan Pan-Montojo, Departamento de Historia Contemporánea, Universidad Autónoma de Madrid
- Professor Roumen Genov, University of Sofia

Physics

Physics programmes:

- Vienna University of Technology, Austria
- Warsaw University, Poland
- Paul Sabatier University, Toulouse, France
- University of Rome La Sapienza, Italy
- Copenhagen University, Denmark

Expert panel:

- Chairman: Professor David W. Hughes, University of Sheffield, Department of Physics and Astronomy.
- Vice-Chairman: Professor Richard Thompson, Imperial College London, Department of Physics.
- Professor Christoph Bargholtz, Stockholm University, Department of Physics.
- Professor of Physics Faculty Vilnius University Gintaras Dikcius, Vilniaus Universitetas.

- Prof. Dr. Ramon Pascual, Universitat Autònoma de Barcelona, Department de Física.
- Director of the Institute of Physics Education, Clemens L.M. Pouw, University of Twente, Department of Applied Physics.
- Professor Peter U. Sauer, University of Hanover, Institute for Theoretical Physics.
- Dr. Frank Witte, Manager of the Master's programmes, Department of Physics and Astronomy of Utrecht University.

Veterinary Science

Veterinary Science programmes:

- Universitat Autònoma de Barcelona, Catalonia (Spain)
- Szent István University of Budapest, Hungary
- University of Glasgow, Scotland (UK)
- University of Ljubljana, Slovenia

Expert panel:

- Chairman: Prof. Dr. Patrick Benard, Ministère de l'agriculture, de l'alimentation, de la pêche et des affaires rurales
- Vice-Chairman: Prof. Dr. Professor Dr. André Parodi (Ecole Nationale Vétérinaire d'Alfort),
- Dr. Francis Anthony BVMS MRCVS (Former president of Federation of Veterinarians of Europe)
- Professor Dr. Jaroslav Hanák (University of Veterinary and Pharmaceutical Sciences Brno, Czech Republic)
- Professor Dr. Jürgen Gropp (Institut für Tierernährung, Ernährungsschäden und Diäteik Veterinärmedizinische Fakultät Leipzig)
- Professor Dr. Henriette Strom (Royal Veterinary and Agricultural University)

In each site visit a **young graduate in Veterinary Science** participated:

- Mr. Simon Doherty BVMS MRCVS (Veterinary practitioner) (Glasgow)
- Mr. Xavier de Paz (Veterinary of B&M) (Barcelona)
- Ms. Katarina Sveticic (LEK International Pharmaceutical group) (Ljubljana)
- Ms. Noémi Szilagyí (Veterinary of CEVA Sante Animale) (Budapest)

Appendix B

List of the members of the Project Planning Group, Management Group and Project Group

PROJECT PLANNING GROUP

ENQA Steering Group:	Christian Thune (PPG Chairman) András Róna-Tas
European University Association:	Leslie Wilson Andrée Sursock (deputy representative)
National Unions of Students in Europe:	Martina Vukasovic Mads Aspelin (deputy representative)
European Commission: Secretaries:	Peter van der Hijden Kimmo Hämäläinen (ENQA Secretariat) (Karl Holm, Finnish Higher Education Evaluation Council, June 2002–September 2002)

MANAGEMENT GROUP

ENQA Steering Group:	Christian Thune (MG Chairman) Ton Vroeijenstijn
Project agencies:	Nick Harris (Quality Assurance Agency for Higher Education, UK) Dorte Kristoffersen (Danish Evaluation Institute) Gemma Rauret (Agencia per la Qualitat del Sistema Universitaria Catalunya)
Socrates/Erasmus Thematic Networks:	Katherine Isaacs (CLIOHNET – European History Network) Hendrik Ferdinande (EUPEN – European Physics Education Network) Tito Horacio Fernandes (EAEVE – Interaction and Co-operation in European Veterinary education).
European Commission: Secretaries:	Peter van der Hijden Tine Holm (Danish Evaluation Institute) Kimmo Hämäläinen (ENQA Secretariat) (Karl Holm, Finnish Higher Education Evaluation Council, June 2002–September 2002)

PROJECT GROUP

Project agencies:	Nick Harris (Quality Assurance Agency for Higher Education, UK) Fiona Crozier (Quality Assurance Agency for Higher Education, UK) Dorte Kristoffersen (Danish Evaluation Institute) Josef Grifoll (Agencia per la Qualitat del Sistema Universitaria Catalunya)
Secretary:	Tine Holm (Danish Evaluation Institute)