

International Comparative Evaluation

Methodological experiences and future perspectives

2003

International Comparative Evaluation

Methodological experiences and future perspectives

2003

Contents

International Comparative Evaluation

© 2003 The Danish
Evaluation Institute
Printed by Vester Kopi

Copying allowed only
with source reference

This publication can be
ordered from:

The Danish Evaluation
Institute
Østbanegade 55.3
DK-2100 Copenhagen Ø

T +45 35 55 01 01
F +45 35 55 10 11

E eva@eva.dk
H www.eva.dk

1	Introduction	3
1.1	Background and purpose	3
2	Lessons Learned	4
2.1	Fulfilment of the purpose	4
2.2	Main conclusions and recommendations	4
3	Methodological Considerations	6
3.1	Focus	6
3.1.1	Selection of focus areas	6
3.1.2	Selection of level	6
3.1.3	Selection of institutions	7
3.2	Organisation	8
3.3	Criteria based approach	8
3.3.1	Frame of Reference	8
3.3.2	Criteria formulation and application	9
3.4	Self-assessment	10
3.4.1	Self-assessment guide	10
3.4.2	Self-assessment process	10
3.5	Site visits	10
3.5.1	Organisation of the visits	10
3.5.2	Interview guides	11
3.5.3	Report and Seminar	11

This report presents the method applied in an international comparative evaluation conducted by the Danish Evaluation Institute (EVA) and the lessons learned from the application of the method. The report is based on EVA's observations and the feedback given by the expert panel appointed for the evaluation and the programmes included in the evaluation. The evaluation covered agricultural science related BSc programmes offered at the Royal Veterinary and Agricultural University, Denmark, University College Dublin, Ireland, University of Hohenheim, Germany and Wageningen University, The Netherlands.

This chapter provides a short description of the background of the initiation of the evaluation and the purpose of the evaluation. Chapter two presents the main methodological conclusions of the evaluation and EVA's recommendations to others who plan to conduct international comparative evaluations. Finally in chapter three the methodology applied in the evaluation is described and assessed.

The programme related results of the evaluation as well as a more comprehensive description and assessment of the method applied are presented in the evaluation report "Agricultural Science – International Comparative Evaluation of Agricultural Science related BSc Programmes". The report can be found at www.eva.dk or bought in a printed version at www.netboghandel.dk.

1.1 Background and purpose

EVA's initiation of an international comparative evaluation has primarily been motivated by recent political developments within higher education taking place at European level. The European perspective on the quality of higher education has, since 1999, been strongly influenced by the process of follow-up to the Bologna declaration of that year. The six objectives of the Bologna declaration and the follow-up process emphasise the need for more comparability and transparency of quality within higher education.

The initiation of the evaluation has been a response to these general objectives and not least the specific objective of promoting European cooperation in quality assurance with a view to developing comparable criteria and methodologies. The view of EVA is that the identification of credible methodologies and procedures for international comparative evaluations is an important step in this direction. EVA is also convinced that there is an obvious need to initiate evaluation projects that seriously set out to try and implement the ambitions of the Bologna process in terms of transparency and comparability of qualifications in higher education. Until now, the number of attempts to conduct such evaluations has been limited.

Reflecting these considerations the objectives of the evaluation have been: to develop and test a common methodological framework and common quality criteria for comparative international evaluations within higher education programmes; to establish mechanisms for continuous quality improvement and cooperation between the institutions participating in the evaluation; and, finally, to stimulate discussion between countries about what constitutes good quality within higher education.

This chapter summarises the methodological conclusions and recommendations which EVA considers relevant to take into account in the planning of future international comparative evaluations. The considerations behind the conclusions and recommendations are described in chapter three.

2.1 Fulfilment of the purpose

The evaluation model and the focused approach applied in the evaluation have generally functioned well and have proved to be useful for handling the complexity that international comparative evaluations are inevitably faced with. The application of common criteria has facilitated the intended comparative perspective of the evaluation, provided a transparent and conspicuous basis for the assessment of the programmes included in the evaluation and ensured that the programmes have been assessed on equal grounds. Furthermore, it has provided an opportunity to identify best (better) practices. Although the criteria developed and tested in the evaluation are of course not directly applicable to other evaluations, they do provide a relevant frame of reference for others engaged in international comparative evaluations.

Variations in the way the institutions have expressed satisfaction with the relevance and usefulness of the process of the evaluation and the content of the evaluation report exist. Yet, the conclusion to be drawn from the feedback from the interviews during and after the evaluation process is that the evaluation has been useful and relevant for the further improvement of the programmes included in the evaluation. At the same time the impression is that the evaluation has stimulated further cooperation between the institutions included in the evaluation. Cooperation, which may include discussion about what constitutes good quality within higher education in the field of agricultural science.

2.2 Main conclusions and recommendations

- The inclusion of national experts (i.e. from the same countries as the institutions being evaluated) in an international panel of experts is recommended as it can ensure that the panel possesses the necessary knowledge about the programmes being evaluated and the cultural, organisational and political framework in which they exist.
- Involvement of institutions in central parts of the process of the evaluation is recommended as it may have a positive impact on the level of commitment of the institutions and the quality of the criteria applied.
- Identical self-assessment guides are recommended as this can ensure that the information provided by each of the institutions are presented in a similar and consistent way, facilitating a comparison between them and an identification of best (better) practises.
- The use of strict and institution-specific interview guides is a useful tool for obtaining effective site visits.
- Generic and general (more than content specific) criteria facilitates the use of the criteria in relation to other evaluations of programmes within higher education, but has the adverse effect that they do not provide the framework for a comparative content-specific assessment of programmes.
- Even a thorough discussion of criteria before they are applied in practice cannot ensure that they are fully understandable and consistent. It is thus important to ensure a process of criteria formulation that includes a critical assessment of the structuring, understandability, clarity, precision and consistency of the criteria. To minimise the risk of different interpretations it is re-

commended that the criteria and/or the self-assessment guide are supported by an explanatory document including a glossary and precise definitions and interpretations of key terms.

- A narrow evaluation scope provides for a concise analysis of the programmes being evaluated. A narrow evaluation scope also allows more time to consider in depth than a broader evaluation scope and allows for the provision of focused documentation material, and thus a report that contains strict analysis, assessments and recommendations. Finally a narrow evaluation scope is time saving for all parties involved, not least for those involved in the self-assessment process. The adverse implication of a narrow evaluation scope is, however, that the evaluation can only present a less than complete picture of the qualities of the individual institutions, compared with an evaluation covering a broader range of aspects.
- When focusing on the bachelor level it is important to be aware of the potential differences in the status and history of bachelor programmes in different countries. On the one hand such differences may imply that institutions with “new” bachelor programmes may benefit from the experiences of those, which have implemented the two-cycle system over a longer period. On the other hand differences in the status of bachelor programmes can reduce the degree of comparability between the programmes being evaluated

3 Methodological Considerations

This chapter provides a description of and reflections on the methodological elements included in the evaluation. The chapter focuses on presenting the main methodological choices related to the evaluation and whether these have fulfilled the intentions behind them. The methodological choices referred to in this chapter relate to the evaluation approach and the evaluation process.

The evaluation has included the following elements:

- Identification of relevant focus of the evaluation and programmes to be included
- A team of evaluation officers from EVA
- An external international panel of experts
- A criteria based approach
- Self-assessments conducted by the programmes included in the evaluation
- Site visits at each of the institutions involved in the evaluation
- A public evaluation report
- A follow-up seminar

3.1 Focus

3.1.1 Selection of focus areas

As a pilot project with a strong methodological focus, it was decided to limit the scope of the evaluation by focusing on only a few aspects of the programmes included in the evaluation. The decision to limit the scope was also based on conclusions drawn from previous international evaluations stressing the importance of limiting the number of focus areas when conducting evaluations across different educational cultures.

The aspects in focus in the evaluation were core competencies, quality assurance mechanisms and internationalisation. These focus areas were all selected due to their relevance in light of the Bologna process.

The experiences with the focused approach have generally been very positive and have provided the following opportunities for the evaluation:

- More time to consider in depth than otherwise would have been the case, given a broader evaluation scope
- Provision of focused documentation material, and thus
- Provision of a report that contains strict analysis, assessments and recommendations
- Time savings for all parties involved, not least for those involved in the self-assessment process.

Having commented on the above positive aspects associated with a focused approach, it should also be mentioned that any focus implies that an evaluation can only present a less than complete picture of the qualities of the individual institutions, compared with an evaluation covering a broader range of aspects.

3.1.2 Selection of level

The scope of the evaluation was also limited in terms of degree level as it focused exclusively on the bachelor level of the programmes.

Generally, the focus on one level 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. In accordance with the overall background of the initiation of the evaluation, the focus on one cycle has, in the opinion of EVA, supported the strategic relevance of the evaluation in the light of ongoing European development within higher education.

Two additional and more practically oriented arguments for the selection of the bachelor level were also decisive. Firstly, the fact that the bachelor programme is a basic education was assumed to facilitate agreement on common quality criteria among the involved institutions (in contrast to the master programme, in which there is a large degree of specialisation). Secondly a focus on one level was expected to provide a more in-depth and concise analysis of the programmes than otherwise would have been possible.

Concerning the first issue, the experience obtained from the criteria formulation process clearly proves that the institutions were very quickly able to reach an agreement on the draft criteria due to the general (and basic) nature of the bachelor level programmes. Viewed in this light, the focus on the bachelor level was a correct strategic choice.

Concerning the second issue, the focus on one level has, according to EVA's best estimate, provided a concise analysis of the programmes involved. This is, however, not only due to the focus on one level, but relates to the overall focused approach applied in the evaluation.

The focus on the bachelor level has not only provided positive experiences, but has also had an adverse impact on the comparative analysis. This especially refers to the different status of the implementation of the bachelor level in the four countries involved in the evaluation. This situation has reduced the degree of comparability between the four institutions. Although this is critical viewed from a methodological perspective, it could be argued that this approach has been beneficial for some of the institutions in the sense that those with "new" bachelor programmes could benefit from the experiences of those, which have implemented the two-cycle system over a longer period.

3.1.3 Selection of institutions

In order to ensure a proper and reliable comparative analysis, EVA decided that at least three, but preferably four, institutions should be included in the evaluation. In addition to an overall desire to obtain a certain geographical coverage, other requirements formed the basis for the selection of institutions. These were:

- The institutions should have a record of commitment to the internationalisation of higher education.
- The institutions should be expected to be motivated to participate in the evaluation and, accordingly, to allocate the necessary time and human resources involved primarily in the self-assessment process and in the follow-up to the evaluation.
- The institutions should be able to appoint representatives from all relevant groups of stakeholders who are able and willing to communicate in English.

These requirements have appeared to be relevant and sufficient for the selection of institutions and throughout the process of the evaluation, the institutions have showed commitment to participate in the evaluation.

Nevertheless, it can be questioned whether the evaluation reflects a "true" European perspective or only reflects the perspective of northern European higher education institutions. The inclusion of southern European universities would, on the other hand, have required a greater focus on the different economic conditions and cultural traditions that influence the work of the universities being evaluated.

3.2 Organisation

In accordance with EVA's usual procedure, a team of evaluation officers from EVA was responsible for the practical and methodological planning and implementation of the evaluation, while a panel of experts - in this case international experts - was responsible for the academic quality of the evaluation, including the assessments, conclusions and recommendations presented in the evaluation report.

The individual experts were identified by asking the institutions to propose relevant national experts from the field of agricultural science who were independent of all the institutions participating in the evaluation. Based on the proposals received from the institutions, EVA approached and appointed those experts who appeared suitable for inclusion in the panel, taking account its desired composition. In addition to including experts with different qualifications within agricultural science, EVA also considered it relevant to include a representative from a typical employer of graduates in agricultural science. The chairperson of the panel was chosen exclusively by EVA based on the primary criteria that this expert had to be independent of the institutions involved in the evaluation and be from another country than those in which the programmes included in the evaluation are based. In view of the strong focus on methodological development, EVA also found it relevant to appoint a chairperson with extensive expertise within evaluation and quality assurance methods.

Providing the institutions with the opportunity to submit proposals for relevant experts served the purpose of ensuring that the panel comprised experts who were respected and considered relevant by the institutions. The decision to include national experts, rather than experts from other countries than those in which the programmes are offered, reflected a belief that this would be the best way to ensure that the panel of experts possessed sufficient knowledge about the programmes being evaluated and the cultural, organisational and political frameworks in which they exist.

Experiences from previous international evaluations stress a number of advantages and disadvantages of appointing national experts, i.e. from the same countries as the institutions being evaluated. The experience of this evaluation is that the inclusion of national experts has been valuable, as it has in fact ensured that the panel of experts possessed the expected knowledge.

At the same time, the choice to appoint a chairperson who was not only independent of the institutions being evaluated, but also from another country to these, ensured among other things a fully independent assessment. Finally, the fact that the chairperson possesses an extensive expertise within evaluation and quality assurance methods ensured that issues important for the achievement of the methodological aim of the evaluation were also continuously addressed during the process of the evaluation.

3.3 Criteria based approach

In national evaluations of educational programmes in Denmark, as well as elsewhere, quality is often assessed in terms of the extent to which the individual programmes achieve their own goals and comply with the legal regulations under which they operate. An approach commonly referred to as the "fitness for purpose" approach.

The goals of the programmes participating in the international evaluation, and the legal frameworks under which they operate, differ. Consequently, the use of the traditional fitness for purpose approach for each programme would not have enabled the intended comparative assessment of how the programmes fulfil common, identical goals. To ensure the comparative dimension, the application of pre-defined criteria was required in order to establish a common framework.

3.3.1 Frame of Reference

The criteria were formulated with reference to a number of different sources. The objectives of the Bologna declaration and the agreements reached at the Prague meeting constituted one

important reference point for the formulation of specific criteria. Other important sources for the formulation of criteria were existing international evaluation models using common quality criteria, and the criteria used in recent international comparative evaluations. Finally, the formulation rested also upon the experience and knowledge EVA has gained from the implementation of numerous evaluations of higher education programmes and from the formulation and use of criteria in the assessment of private education programmes.

Although criteria formulation benefited greatly from the many different sources and earlier experiences, it was nevertheless vital to take into account the specific conditions which characterise an international comparative evaluation. Firstly, there are considerable differences between educational cultures, national traditions and regulatory systems within which the individual programmes operate. Secondly, the aim of developing a methodology for international comparative evaluations implied an obligation to ensure that the criteria formulations were sufficiently flexible to allow them to be replicated in other international programme evaluations with a comparative perspective. Thirdly, the variation in programme content represented a significant challenge to the development of commonly relevant criteria for agricultural science programmes that would also provide space for the expression of individual priorities and qualities.

To overcome these obstacles, and to assure a high level of common applicability and relevance, EVA developed a set of criteria requirements which served as the framework for the criteria formulation. The character and content of the draft set of criteria were driven by the following requirements:

- *Breadth*: To ensure the criteria respect specific national traditions, concerns and priorities, and do not hinder diversity, the criteria must be formulated broadly enough to allow for variations.
- *Uniformity*: The criteria should be the same for all the programmes participating in the evaluation. This ensures that the programmes are assessed on an equal basis, that the assessments are transparent and that a comparative perspective is enabled.
- *Reference to level*: In order to operate with one set of criteria, this set has to be formulated with reference to the BSc as a single level, irrespective of the variations in the nominal duration.
- *Precision*: The criteria must be precise enough to allow an assessment of how they are fulfilled by the individual programmes.
- *Internal consistency*: The set of criteria must be internally coherent.
- *Topicality*: The criteria must reflect present objectives and developments within the area of higher education in Europe

3.3.2 Criteria formulation and application

The institutions were included in the process of the formulation of the evaluation criteria. This approach proved itself to be a relevant way to ensure that the criteria were generally relevant and understandable to the institutions. It must, however, be stressed that even a thorough discussion of the criteria before they are applied in practice cannot safeguard against later misinterpretations, or ensure that they are fully understandable and consistent.

As part of the self-assessment process, the institutions were asked to critically assess the quality of the criteria used in the evaluation. Furthermore, the quality of the criteria was discussed at the site visits. Through this it became apparent that in particular the definitions of important terms led to discussions at the individual institutions during the carrying out of self-assessment, despite the fact that they were agreed upon earlier in the process of the evaluation. Similarly, terms that were not supported by definitions led to different interpretations, which in turn had a negative impact on the comparability of the information provided in the self-assessment reports.

These experiences illustrate the importance of ensuring a process of criteria formulation that includes a critical assessment of the structuring, understandability, clarity, precision and consistency of the criteria. To minimise the risk of different interpretations it is recommended

that the criteria and/or the self-assessment guide are supported by an explanatory document including a glossary and precise definitions and interpretations of key terms.

Despite the critical points raised above the application of the common criteria did facilitate the intended comparative perspective of the evaluation, provided a transparent and conspicuous basis for the assessment of the programmes included in the evaluation and ensured that the programmes were assessed on equal grounds. Furthermore, it provided an opportunity to identify best (better) practices. Thus, the conclusion to be reached is that the criteria developed and tested in the evaluation provide a relevant frame of reference for others engaged in international comparative evaluations. At the same time, it must be concluded that there is still a need for substantial improvement of the criteria if they are to form a general frame of reference for international comparative evaluations within the field of higher education.

3.4 Self-assessment

3.4.1 Self-assessment guide

Each institution involved in the evaluation conducted a self-assessment. To facilitate and structure the self-assessment process each institution was provided with an identical self-assessment guide. The purpose of the identical guide was also to ensure that the information provided from each of the participating programmes would be presented in a similar and consistent way, thus facilitating the comparison across the institutions and an identification of best (better) practices. The guide contained a number of questions related to each of the focus areas of the evaluation as well as a number of general questions concerning the programmes.

The fact that the self-assessment guides have been identical has generally ensured that the information provided by each of the institutions has been presented in a similar and consistent way, facilitating a comparison between them and an identification of best (better) practises. The comparability of the information could, however, have been even further increased by the provision of a stricter format for the self-assessment report than just an overall page limit. For instance, a format including predefined space or word limits for the answers to each of the focus areas and ready-made tables for insertion of required quantitative data.

3.4.2 Self-assessment process

Each institution documented the results of its self-assessment in a self-assessment report. As required the self-assessment reports contained both descriptions and assessments of the present status of the programmes under evaluation in relation to the focus areas of the evaluation. The self-assessment groups generally included at least one representative from each of the relevant stakeholder groups at programme level, including management, teaching staff, students and administrative staff.

The institutions were given two and a half months to the self-assessment process and the impression is that this was adequate. Regarding the value of the self-assessment for the institutions their feed-back generally reveals that the self-assessment process has been useful and relevant for the further improvement of the programmes included in the evaluation.

3.5 Site visits

3.5.1 Organisation of the visits

After receiving the self-assessment reports, the panel of experts visited each of the participating institutions. The purpose of the site visits was to provide the panel of experts with an opportunity to ask the institutions to elaborate unclear and less substantiated sections of the self-assessment reports and to validate the information provided in the self-assessment report.

Each visit comprised a number of separate interviews with the different groups of stakeholders who, in one way or another, were engaged with the programme(s) included in the evaluation. The purpose of conducting separate interviews with different stakeholder groups was to

validate the content of the self-assessment reports. In other words, the interviews were used to clarify the opinions, perspectives, etc. of the different stakeholders in relation to the information provided in the self-assessment report.

The organisation of the site visits served the intended purpose of validating the content of the self-assessment reports by providing the panel with a clear picture of the opinions and perspectives etc. of the different stakeholders.

The importance of this is illustrated by that fact that the site visits revealed some remarkable differences in the ways the different stakeholders agreed with descriptions and assessments provided in the self-assessment reports. This fact further illustrates the importance of using the visits as a means of getting a clear picture of how the self-assessment process had worked, including the writing of the report. Great insight was achieved by asking both the self-assessment group about the process, and each of the stakeholders being interviewed about the extent to which they had been involved in the process. This approach provided a clear picture of the extent to which stakeholders agreed with the information provided in the report.

3.5.2 Interview guides

To ensure that the site visits functioned as a useful supplement to the self-assessment reports, institution specific interview guides were prepared and used at the site visits. Accordingly, the content of the guides differed, reflecting the differences in the content and quality of the self-assessment reports.

The use of the relatively strict interview guides, including a number of different questions for each of the groups of stakeholders being interviewed, counteracted the dynamic flow of the interviews somewhat and often meant insufficient time to follow up on answers. On the other hand, the use of the guide, and the fact that it was institution specific, did ensure that the site visits functioned as a useful supplement to the self-assessment reports. This, in turn, facilitated the substantial comparative assessment of the programmes.

3.5.3 Report and Seminar

Despite variations in the self-assessment reports, the process by which they were prepared and the organisation and carrying out of the site visits, the two forms of documentation material complemented each other. This enabled a comprehensive assessment of the quality of the programmes included in the evaluation in terms of the selected focus areas of core competencies, quality assurance mechanisms and internationalisation.

The panel of experts assessed the extent to which each of the programmes complied with the criteria used in the evaluation. Whenever relevant, the recommendations following the assessments included a "best/better practice perspective" in the sense that the panel of experts recommended that the good practice(s) of one or more institution(s) in one area of operation was adopted by the others as a frame of reference when implementing the recommended changes.

The draft report was submitted to the institutions for comments and then finalised for publication in November 2002. Shortly after the publication of the report EVA invited the institutions and the panel of expert to participate in a seminar with focus on how the results of the evaluation could be followed up.

At the seminar the feed-back from the institutions on the recommendations provided in the evaluation report revealed that these were considered useful for the further improvement of the programmes included in the evaluation. In particular, the institutions highlighted the value of focussing on identifying examples of good practise and using the identified good practises as reference points in the formulation of recommendations. Moreover, the institutions agreed that this approach made them very aware of their possibilities for cooperation and for exchanging experiences to further mutual learning and quality improvement.