

Criteria based evaluations

EVA's experience in evaluations based on criteria

Criteria based evaluations

EVA's experience in evaluations based on criteria

2004

Contents

Criteria based evaluations

© 2004 The Danish Evaluation Institute

Copying allowed only with source
reference

The publication is only published on
our website

The Danish Evaluation Institute
Østbanegade 55, 3.
2100 Copenhagen Ø

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

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

ISBN 87-7958-153-6

1	Introduction	5
2	Conclusions	7
3	Clarification of concept	9
3.1	Criteria	9
3.2	Indicators	10
3.3	An example	10
4	Experience	11
4.1	The basis for the criteria	11
4.2	Transparency	11
4.3	Comparability	12
4.4	Involvement	12
4.5	Control and development	12
4.6	Number of criteria	13
4.7	Clarification of concept	13
4.8	Final comments	13

Within the last two years EVA has gathered its first experience from conducting evaluations using explicit criteria as a basis for assessment. First and foremost an evaluation of three master programmes in social science and an international evaluation of the Master of Science (Agriculture) were conducted in four European countries. Both of these evaluations assumed the character of a pilot project in which EVA had attached importance to gathering methodical experience, including, among other things, the strengths, weaknesses and potential that this type of evaluation could offer. EVA has conducted more criteria based evaluations since then: an evaluation of the Danish programmes in classical music and a pan-European evaluation of physics. The latter was completed in 2003. Last but not least EVA has prepared criteria for the professional bachelor programmes commissioned by the Ministry of Education for specifying the expectations of the ministry for the award of the professional bachelor title.

The reason why EVA is interested in gathering methodical experience in criteria based evaluations is especially motivated by international trends where the use of criteria in evaluations has become more widespread in recent years. This is among other things due to the emphasis on the need for openness and transparency as part of the internationalisation of higher education related to the implementation of the Bologna process.

On the basis and in the light of the fact that this type of evaluation is relatively new under the remit of EVA, there is a significant need to gather and systematise preliminary experience in these evaluations.

Aims and contents

The aim of this publication is thus to gather EVA's experience in criteria based evaluations and thereby qualify the use of criteria in future evaluations. Moreover, the publication is intended to clarify EVA's use of the criteria concept and contribute to extensive use of the experience.

The publication will, among other things, illustrate the following issues:

- What is the added value of criteria based evaluations compared to evaluations that do not make use of explicit criteria? Which new elements of methods does this type of evaluation generate in EVA's evaluations, and to what extent is it compatible with a fitness-for-purpose approach, i.e. an approach based on the aims and principles defined by the educational programmes themselves?
- Who sets the criteria and how is embeddedness safeguarded with the central players in the evaluation?
- How do you avoid that criteria based evaluations lead to standardisation and limited new development of the evaluated programmes?
- How is the aspect of development best safeguarded in the evaluations?

Despite the challenges, value has clearly been added in using criteria as the basis of assessments in the pilot evaluations. The following has been especially important:

Openness and transparency

Criteria offer openness and transparency in the evaluation because criteria explain the basis of assessment and added value. It is clear on what terms the evaluation is conducted, and the basis for identification of good quality is open and known. You might say that the criteria constitute the quality concept of the evaluation.

Focus

Criteria provide a good focus in the evaluation and a better opportunity for maintaining internal focus on a current basis compared to the evaluation group and externally compared to the educational establishments that have been evaluated. A focused evaluation demands, however, that in formulation and choice of criteria a comparably harsh prioritisation be made between what is important and what is less important, to keep the number of criteria at a reasonable level. The result of such focus should be that all the criteria have more breadth and that they do not overlap and replicate each other.

Co-operation and dialogue

Criteria constitute a sound and constructive basis for co-operation between EVA and the evaluation groups and for dialogues with the educational establishments. By incorporating the evaluation groups and the educational establishments in the formulation and selection of criteria, highly academic precision, experience of co-ownership and thus a higher degree of legitimacy can be ensured in the evaluation. Moreover, EVA's special experience and competencies in formulating criteria make it important that the institute maintains its position as a guarantee for the solidity of the final criteria.

Integration of standards

Criteria constitute a sound basis ensuring that the evaluation includes and balances all relevant standards. Experience shows that the criteria may include statutory requirements, more qualitative aspects such as the requirements and expectations of the evaluation group and the aims and principles of the educational establishments themselves, for example.

Control and development

Criteria provide ample opportunity for combining control and development factors in the evaluation by using a mix of different types of criteria, such as minimum and best practices criteria, fitness-for-purpose based criteria and open and closed criteria, for example.

Reporting

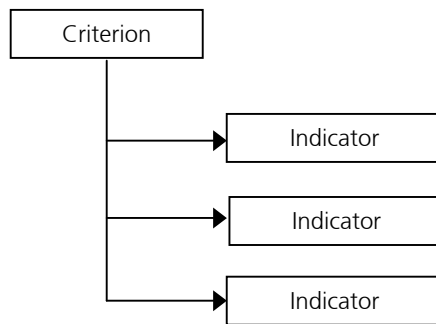
Criteria enable relatively brief reporting, where focus can be put on criteria that have not been achieved.

Moreover, there is a challenge in balancing the inclination to focus on the individual criteria at the expense of the holistic perspective.

3 Clarification of concept

There is no single national or international definition of the meaning of criteria as a concept in evaluation, and it is often confused with concepts such as standards, benchmarks and indicators. Criteria are especially often used as a synonym to standards. In order to clarify the meaning of the individual concepts and their inter-correlation, this chapter suggests some definitions. The definitions are aimed at capturing the application of the concepts rather than giving a more theoretically embedded definition. This means among other things that the definition of standards and criteria, respectively, are not mutually exclusive, but on the other hand it identifies various functions, levels of details etc. that are relevant for practical application and exercise of the concepts.

As can be seen, EVA envisages a hierarchical correlation between the concepts where the criteria are of a more general character than indicators. The correlation between the concepts is as follows:



3.1 Criteria

“Criteria are operational requirements for quality”.

EVA defines *criteria* as a term for operational statements on quality in teaching and education. Criteria are thus focused requirements for quality that are close to the operational level though they do not necessarily contain a specific scale.

As regards EVA’s evaluations it will mean that a number of explicit requirements can be formulated for evaluations which the programme or educational establishment that has been evaluated must comply with in order to be accepted as being of good quality. In contrast to EVA’s traditional fitness-for-purpose evaluations, criteria based evaluations include pre-defined statements for acceptable or extremely good quality.

Fitness-for-purpose and criteria

Fitness-for-purpose is a term for evaluations where the basis for assessment of the quality is embedded in the aims and principles of the educational programme itself. However, there is not necessarily a conflict between fitness-for-purpose and the use of criteria as a basis for evaluation. A number of EVA’s criteria based evaluations have included elements of fitness-for-purpose in that the criteria were very accommodating and open to interpretation from the individual establishments. This has been the case, for example, when the criteria were attached to the aims and objectives of the said establishments or programmes.

Criteria as a basis for assessment

However, any criterion must express quality precisely to ensure that a transparent basis for assessment be established. Criteria are thus testable statements that are formulated on the basis of accepted statements to examine whether the required quality level is achieved. There can and should be close correlation between criteria and statements.

Minimum requirements and best practice

Criteria can express both minimum requirements and best practice. "Minimum requirements" means criteria that describe conditions that shall be satisfied in order for the quality to be acceptable – typically statutory conditions. "Best practice" means requirements that must be fulfilled for the educational activity to be acceptable as being of extremely good quality level within a specific area.

Benchmarks

Benchmarks are a particular type of criteria. Benchmarks are criteria that are embedded compared to the practice of others. The use of explicit references and indicators in relation to systematic comparisons that are applied to identify potential improvement or improved efficiency of the educational programmes or establishments involved. Benchmarks are always embedded in analyses of best practice or extremely good quality in order to make the potential improvement of the educational programmes or establishments involved visible.

3.2 Indicators

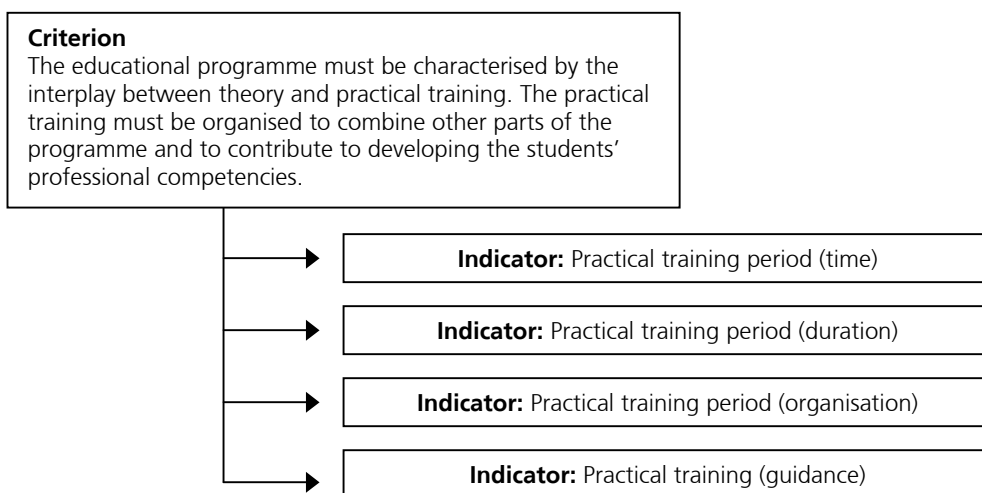
"Indicators are concrete objectives for criteria".

EVA defines *indicators* as concrete, measurable, qualitative or quantitative variables that express central characteristics of the establishment, programme or teaching required to be evaluated by the use of criteria. The indicators should altogether give a comprehensive picture of whether the programme is in compliance with the criteria applied.

Generally a number of indicators are applied to present a satisfactory identification of a criterion. Setting up indicators is conducted through operationalisation of the criteria in a way that a high degree of validity is safeguarded through close correlation between indicators and criteria. Indicators are related to all aspects of an educational programme and they may present direct and indirect descriptions.

3.3 An example

In order to make the concepts more clear, an example of how criteria and indicators can be formulated in practice is given below.



As shown all the indicators provide the basis for information that is necessary to be able to assess whether the requirements of the criterion are fulfilled.

The evaluations that this paper is based upon include a number of similarities that reflect the current exchange of experience that has taken place between the projects. However, the examples also illustrate a considerable variation in approach and contents. This variation shows that EVA has aimed at testing different models and approach. Differences should also be included in EVA's criteria based evaluations in the future, having regard to the different aims, objectives and focus of the evaluations.

This chapter sums up EVA's most important experience from criteria based pilot evaluations.

4.1 The basis for the criteria

The criteria in the pilot evaluations stem from a wide range of sources. In the evaluation of the three master programmes in social science the criteria were formulated on the basis of relevant legislation and existing surveys in this area, experience from the Evaluation Institute and the academic knowledge and experience of the evaluation group and the educational establishments. The criteria that were applied in the evaluation of the Master of Science (Agriculture) programme are based on a similar basis, though the evaluation group did not play the same role. Moreover, the legislative framework was not of similar importance due to the international character of the evaluation and differences in national laws in this area. The same applied to the evaluation of the Danish programmes in classical music which was primarily inspired by other criteria based evaluations, studies of literature and articles and the academic knowledge of the evaluation group in this area.

The criteria for the professional bachelor programmes are in general legally based on the ministerial order on professional bachelor programmes and the Ministry of Education has assessed the judicial validity of the criteria.

In all of the pilot evaluations, the requirements below were, moreover, applied as guidelines for the formulation of criteria. The criteria of the evaluations must:

- be relevant for the educational level in question
- be uniform and have broad validity
- be accommodating and include an element of fitness-for-purpose
- be operational and form the basis for real assessment
- focus on implementation and not only principles
- be in consistency with each other.

4.2 Transparency

Experience from all of the pilot projects indicates that criteria used as a method for evaluation has a particular strength by providing transparency in the basis of the assessment. For the pilot evaluations it has had the effect that the evaluation group has had a common denominator for the perception of "good practice". This especially applies to the evaluation of the three master programmes in social science and the evaluation of the Danish programmes in classical music where the evaluation group were involved in setting up the criteria and where the criteria thus made the evaluation group's interpretation of criteria visible. The criteria clearly expressed what the evaluation group considered good quality.

For the educational establishments the use of criteria has made clear the premises that were established for the assessment of their strengths and weaknesses and recommendations of the evaluation. For EVA the use of criteria has turned out to be an effective tool for retaining relevant focus in the evaluations.

4.3 Comparability

Moreover, as a method the criteria has the advantage that they are appropriate as a focal point for systematic comparisons. This has especially been the case in the evaluation of the Danish programmes in classical music and the Master of Science (Agriculture) programmes where the quality of the programmes was compared across educational establishments and borders. Moreover, there seems to be other factors such as integration of the element of fitness-for-purpose, for example, that can make comparisons difficult.

4.4 Involvement

Experience drawn from the pilot evaluations shows that both the project teams and evaluation groups play a central role in connection with the formulation of the criteria.

The evaluation group participated in the formulation of the criteria both for the evaluation of the three master programmes in social science and the evaluation of the Danish programmes in classical music. Throughout the evaluation process, the criteria put the members of the evaluation group under an obligation vis-à-vis the focus and criteria set up and they were appropriate for structuring the discussions of various aspects in the programmes.

The dialogue with the evaluation groups about the criteria is important due to the academic qualification and to safeguard co-ownership, but EVA also possessed special experience and competencies for formulating criteria. It is therefore important to ensure that EVA retains its position as a warranty for the solidity of the final criteria.

The educational establishments have been involved in the formulation of the criteria of the three evaluations to a varying degree. Moreover, experience from the pilot evaluations shows that involvement on the part of the establishments has contributed to legitimacy and co-ownership in the external assessments and has safeguarded a safe starting point for the further process.

Experience drawn from the evaluation of the Master of Science (Agriculture) programme was different in that respect. Similar to the evaluation of the three master programmes in social science the experience drawn from here showed that involvement on the part of the establishments to some extent contributed to safeguarding that the criteria have been complete and relevant for the educational programmes evaluated. Moreover, involvement meant that the criteria were close to practice and did not only express statements of intent and principles. However, involvement of the establishments must be balanced accordingly to safeguard an independent basis for the evaluations.

4.5 Control and development

General experience derived from the four projects shows that the perspective for improvement is best safeguarded by including open criteria rather than closed ones because they suggest reflection in the programmes and the establishments. However, this should rather be seen as a choice between minimum criteria and criteria for best practice. When the criteria are attached to a best practice perspective, it clearly helps focussing on opportunities and needs for quality improvement. An open minimum criterion may thus suggest reflection on different ways in which to comply with the criteria, but it will typically not have as far-reaching perspectives for improvement as best practice criteria. In this context, however, it is important to pay attention to the fact that the criteria must relate to "genuine" examples, which means that it will not be unrealistic to meet the criteria and that the establishments are not measured on abstract aims and objectives. In this context it is worth considering the use of good practice instead of best practice since the best establishments or educational programmes can achieve results that are not optimal in an absolute sense. Minimum criteria are appropriate to use to the extent that you wish to apply the criteria to express requirements and conditions in relation to relevant legislation in education.

The four projects have all applied a number of criteria that are fitness-for-purpose based. Experience in integrating the element of fitness-for-purpose has shown that it is suitable to safeguard that the criteria do not become uniform and do not restrict further development of the programmes. Moreover, the element of fitness-for-purpose helps safeguarding the accommodation in the assessments that EVA is striving for. However, the element may have the effect that the programmes cannot be compared as such due to different aims and objectives in the programmes. This is especially a problem for evaluations with strong comparative objectives. It is difficult to point out which of the two factors carries the greatest weight. It depends on the context, amongst others the aim of the evaluation.

4.6 Number of criteria

The pilot evaluations have shown that some central challenges exist in terms of balancing the criteria based evaluations as regards the number of criteria and the complexity of the individual criteria, i.e. how many dimensions the individual criteria cover.

General experience shows that the number of criteria was quite large, and therefore overlapping and repetitions appeared. In retrospective the criteria could have been reorganised, which could have helped avoiding repetitions and made it possible to distinguish between the important and the less important. The number of criteria was also pointed out by the educational establishments as requiring an (unnecessary) number of resources compared to the self-evaluation process.

In the evaluation of the Danish programmes in classical music and the Master of Science (Agriculture) programmes it was also experienced that in practice it has been difficult to distinguish between the individual criteria. This applies especially to the criteria for aims and contents, respectively.

In connection with the evaluation of the three master programmes in social science it was emphasised that some of the criteria were too complex and included too complex relations. This had the effect that the dynamics of the criteria to some extent were diminished. On the other hand, it should be emphasised that it is important to ensure that the criteria do not become so simple that they end up expressing a too simple quality concept and thus fails to correspond to the complexity of the educational factors that are evaluated.

4.7 Clarification of concept

An important experience drawn from the evaluation of the Master of Science (Agriculture) programme was the importance of making accurate definitions of the central concepts that appear in the criteria, e.g. "professional qualifications". By doing so the risk of misunderstandings and misinterpretations in the self-evaluations reports prepared by the educational establishments can be avoided. This is especially important in an international evaluation where the self-evaluating establishments in many cases operate in the English language. Different interpretations of the concepts may lead to different responses that can be difficult to compare across nations. In future evaluations it should therefore be considered that a "glossary" explaining terms be prepared and a copy enclosed with the self-evaluation guide.

4.8 Final comments

The initial experience that EVA has drawn from criteria based evaluations has been of great value to the new criteria based projects that the institute has recently introduced and that it will consider initiating in the future. In parallel with EVA's own experience and reflections, the institute will continuously follow the experience drawn by other evaluation organisations in this area.