

Policy Analysis

Subject Assessments for Academic Quality in Denmark: A review of purposes, processes and outcomes

Bjørn Stensaker
NIFU

Executive summary

With a dual purpose related to accountability and improvement, Denmark has had a system of external quality monitoring since 1992. Organised through an independent agency (EVA), systematic evaluation of study programs has been the dominant method for quality assurance for a number of years. By the use of data and observer triangulation, stability in procedures and through extensive dialogue, the evaluations have resulted in noticeable changes in teaching and learning, study program objectives, and triggered dialogue and reflection both within higher education institutions as well as between higher education and its stakeholders. In 1999, the agency responsible for the study program evaluations was made permanent by an act in parliament, and has since then expanded into evaluating primary and secondary education in addition to higher education. At present, the greatest challenge for the Danish national evaluation system is to adjust a well-functioning domestic system of study program evaluation to the emerging international trend related to accreditation and convergence in how quality should be assured in the international marketplace.

Introduction

Higher education in Denmark is mainly public and consists of about 110 institutions including traditional universities, vocationally oriented colleges, and more specialised higher education institutions (in art, agriculture etc.). The Ministry of Education approves all public higher education institutions. Private institutions may operate without governmental approval, but then run the risk that their students will not be eligible for the state student grant. The higher education system can be divided into a university sector and a college sector (a binary system). The university sector consists of 12 institutions, and the remaining institutions all belong to the college sector. The substantial number of (small) colleges triggered the government in the late 1990s to stimulate voluntary amalgamations in this sector (Gornitzka et al 2001: 16). This process is still ongoing and the first amalgamated “Centres for higher education” (CVU) have been established.

The degree system has three levels: bachelor studies (3 years), master degree studies (5 years), and the PhD-degree (additional 3 years). However, within the college sector one can find study programs that deviate from this structure, and that rather could be described as “short cycle” (1-3 years), “medium-cycle” (3-4 years), and “long-cycle” programs (5-6 years) (Thune 2001: 3). As such, the degree system is rather complex with limitations related to the possibility of transferring credit points within the system, especially between the college and the university sector. Denmark has a system of external examiners which partly comprises teachers/professors from other institutions, and partly labour market representatives. The role of the external examiners is to assure that students are treated fairly and to assure an equivalent national level of assessment across schools and institutions (Kristoffersen 2003: 26).

As in other OECD-countries, Denmark has during the last ten to fifteen years experienced a rapid increase in student numbers. In the recent years, the gross intake to higher education has been between fifty and sixty percent of the relevant age group (Thune 2001: 3). Most of these students enrol on long-cycle higher education programs. Higher education institutions are responsible for admissions, but admission requirements are set by the Ministry of Education. In some programs, for example in Medicine, the Ministry still sets the admission number. In general, student numbers in study programs vary according to student preferences and choice. There are no tuition fees in the public sector.

The steering and funding of Danish higher education have changed considerably during the 1990s. The trend has been to delegate more responsibility from the Ministry of Education to higher education institutions. One may claim that the changes in the steering of the sector have stimulated the autonomy of the institutions, even though the power and autonomy of Danish universities have been historically quite strong. However, strategic behaviour and strong institutional leadership have not been a central characteristic of Danish universities. Hence, in 2000 the Ministry of Education launched what may be termed as “development contracts” between the Ministry and the individual institution. The purpose is to agree on more long-term objectives and targets (four year periods) and to enable the institutions to market themselves better. It is voluntary for the institutions to join in the contract arrangements, and so far no sanctions or rewards have been linked to these instruments.

The changes in the steering of higher education have been followed by a change in the funding of higher education with more emphasis on lump-sum allocations and output measures (Gornitzka et al 2001: 19). This means that the higher education institutions can decide on how to allocate resources internally. The most important output measure (the “taximeter-system”) is a combination of different indicators related to student numbers, the cost of studies in different disciplines and subject fields, and the number of credit points and exams taken. Research is funded separately. Four streams of money comprise most of the research funding: a lump-sum from the Ministry, allocations from different domestic research councils, applied research programs, and some funds from the Danish fund for basic research (DGF).

The described changes in the steering and funding of higher education in Denmark in the last ten to fifteen years have also had implications for how academic quality assurance is conducted. Traditionally the country had a decentralised system of quality assurance, which left quality assurance up to the individual institution, with the external examiner system as the key component. In 1992 the Ministry of Education established the Danish Centre for Quality Assurance and Evaluation of Higher Education (EVA¹), and instructed the centre to *conduct systematic evaluation of all study programs offered in higher education* within a seven-year period. Hence this centre can be interpreted as a more centralised and independent actor in the field of academic quality assurance. Why the political authorities at that time perceived a need for systematic evaluation at the national level is discussed below.

The policy problem

In the spring of 1992 a majority of the parties in the Danish Parliament arrived at a number of compromises on higher education, which led in the following year to a reform of the entire educational system. The stated objectives of the reform were to ensure (Thune et al 1996: 21):

- a higher degree of institutional freedom and autonomy combined with a tightening of each institution's management structure,
- a better balance between supply of and demand for study places,
- the quality of the study programs according to international standards.

The reform implied a new study structure (the bachelor/master/PhD-system), a new Act on universities, which reorganized the political and managerial governance of the institutions (reducing the number of democratically elected governing bodies and introducing external representation in the academic senate and in faculty boards), an introduction of the taximeter-principle (an output-based funding system), and the establishment of a national system for the evaluation of higher education (conducted by a newly established agency for conducting such evaluations [EVA]).

When looking at the stated objectives of the reform, the background for the reform also comes to the fore. First, a huge increase in the number of students that applied for higher education. Second, Denmark faced at that time constraints on public spending, which triggered a focus on the efficiency and the effectiveness of higher education. Third, worries that an expansion of higher education could lead to a lowering of the academic quality. Fourth, the international commitments that Denmark had in relation to the European Union and their student exchange system (Erasmus).

The establishment of a national system for evaluation and an independent agency for carrying out such evaluations are in various ways related to the drivers behind the reform. The establishment of the EVA agency could be described along a number of different perspectives:

- a) The creation of the system of study program evaluations could be interpreted as a governmental response to perceived needs for more *efficiency and output-orientation* in Danish higher education. The share of resources spent on higher education, due to the increasing number of students among other reasons, triggered a need to check how resources are spent and to identify "organisational slack" inside higher education institutions. The systematic evaluation of all study programs offered in Danish higher education can be seen as an indicator of such an orientation.

¹ The abbreviation EVA is used throughout the document even though the organisation "Centre for Quality Assurance and Evaluation" (Evalueringsscenteret - EVC) changed its name and formal status in 1999 to the "Danish Evaluation Institute" (Danmarks Evalueringsinstitut - EVA).

- b) At the same time, study program evaluations could also be seen as an attempt to balance the *centralisation-decentralisation dilemma* in Danish higher education. While major parts of the 1992 reform intended to give institutions more autonomy, the establishment of a national evaluation system could be interpreted as a form of centralisation attempt when it comes to the quality assurance. Thus, the evaluations in this perspective represent the need to maintain control even in a more decentralised system.
- c) Since the evaluations were established with a double purpose of accountability and improvement, it is also possible to see the establishment of the evaluation system in a more *developmental* perspective. The decentralisation of authority and responsibility to institutions meant that the institutional leadership had to take on a stronger and more strategic role. However, this is a role that breaks with the traditions of institutional leadership in Danish universities. The traditional power structure in higher education centred round the departments and disciplines (Gulddahl Rasmussen 1997, Foss-Hansen 1997). The national evaluation system could in this perspective be interpreted as being a “support structure” for the institutional leadership (see also Stensaker 1999: 257-258).
- d) The notions “knowledge society” and “knowledge economy” and the role of higher education and research in these developments have had a powerful influence in the political debate on higher education in the last two decades. One of the important elements of the knowledge society is that higher education needs to establish better links with the world of work (Rasmussen 1997). In the new evaluation system, these links are very visible. Not only are members of industry and society part of the review panels, but graduated students are also, after a few years at work, asked about the relevance of their study program in relation to their current job. In this perspective, the study program evaluations could be seen as an instrument for increasing the *relevance* of higher education for the society and the world of work.
- e) Finally, one could also interpret the establishment of study program evaluations as a form of political accountability. Not only higher education, but also those responsible for higher education at the political level need to be accountable to the larger society. Hence, the creation of a national system for evaluation, and an agency responsible for carrying out such tasks could be interpreted as being an important *symbolic* action, by which politicians can show the public that something “is done to assure quality”. One indication of this is that how evaluations should be followed up was almost a non-issue in Denmark in the first few years of the 1990s (Askling et al 1998: 9). Not the outcomes, but the fact that evaluations were conducted seemed, in other words, to be the important thing.

To pinpoint the policy problem in accurate terms is, in other words, somewhat problematic. However, the five perspectives mentioned above cover most of the arguments related to the establishment of the national evaluations in 1992, and can be said to have represented a formidable challenge for the leadership of EVA to balance these various needs and expectations.

Content of the policy instrument

The mandate for EVA, provided by the Ministry, instructed the centre that future evaluations had to focus on the study program level, that both control (accountability) and institutional improvement had to be a part of any procedures launched, and that evaluations were not a voluntary activity for the institutions. However, the results of the evaluations were not linked to funding (Evalueringscenteret 1998: 16-17). The evaluation system was not created on a permanent basis, but was set up for an initial period of seven years, and on the condition that the system and EVA itself should be subject to an evaluation when deciding whether evaluations should become a permanent activity. The political focus on study programs can probably be related to the huge number of small

higher education institutions in Denmark, and the fear that the institutions could not be trusted as assurers of quality (see also Thune et al 1996). Also, the systematisation meant that all study programs were treated equally – a particular feature in the Scandinavian culture (Smeby 1996). EVA was created as an independent body. (See Box 1 for EVA’s legal and organizational framework.) The Ministry of Education was not to instruct the centre, but the National Educational Councils (NEC) (in humanities, science, social sciences etc.) were given the right to decide the chronological order of the evaluations, and thus could be seen as the bodies responsible for the initiation of a given evaluation.

Box 1. EVA’s legal and organizational framework

1. Legal framework

Two legal documents regulate EVA’s activities. The most important one is the Danish Evaluation Institute Act. The Ministry of Education has established a set of regulations for EVA that specifies the act. The regulations are as legally binding for EVA as the parliamentary act, but it is within the authority of the Minister of Education to amend the regulations within the framework of the parliamentary act.

The legal framework regulates the relationship to the Ministry of Education and specifies:

- EVA’s right to initiate evaluations;
- the governance of the agency;
- the distribution of responsibilities with regard to evaluation;
- core methodological principles.

2. Main stakeholders

Within the field of higher education, the Ministry of Education and the new Ministry of Science, Technology and Development (established after the Danish election November 2001) represent the main stakeholders, e.g. they have to approve the annual plan of action and the budget. Besides these formal relations, EVA has regular contact-meetings with the Ministry of Education and is in the process of establishing a network at staff level. In addition to the ministries, EVA has maintained contact with stakeholders from the higher education community. EVA meets with the Danish Rectors’ Conference, which represents all universities in Denmark, and EVA’s Committee of Representatives, which comprises members from different sectors of the education system.

3. Governance

EVA is an independent institution formed under the auspices of the Danish Ministry of Education. Eva is governed by a board. The Board is responsible for the overall supervision of the Institute, including the annual action plan, and appoints the management of the Institute. The appointment of the Executive Director must be formally approved by the Minister of Education. The Executive Director manages EVA and is responsible to the Board. The Board formally approves the appointment of other staff.

The Board consists of 10 members and a chairman. The Danish Minister of Education nominates the Chairman. The 10 members are appointed by the Minister upon the recommendation of the Ministry’s advisory boards. Thus, the Board does not automatically include representatives of the higher education institutions. The Board is appointed for a three-year period with the possibility of reappointment.

In addition to the Board, a Committee of Representatives is established as a mandatory part of EVA’s organisational set-up. The Committee of Representatives comments on EVA’s annual plan of action, the annual report and the priority of planned activities. The Committee comprises 27 members. They are appointed by organisations from the following sectors: school proprietors, school associations, school boards and employers; rector’s conferences and school managers; management and labour organisations; teachers’ organisations and students and pupils bodies. In addition, the Committee of Representatives itself appoints two experts with international evaluation experience.

The Board draws up the programme for the next year’s activities based on the recommendations of the Executive Director. The Minister of Education approves the annual plan of action. In addition to the evaluations conducted on its own initiative, EVA may conduct evaluations on the request of authorities responsible for education.

Source: EVA 2002.

EVA's formal mandate was (Thune 2001: 7):

- to initiate systematic evaluation of all study programs in higher education in Denmark including the university as well as the non-university sector (within a six year period),
- to develop appropriate methods of the recognition of study programs making them eligible for governments funding and student loans,
- to inspire and guide the institutions of higher education in aspects concerning evaluation and quality,
- to compile national and international experience on evaluation of the educational system and quality development.

Given the many interests surrounding the establishment of the study program evaluations, it is perhaps not surprising that active dialogue and consultation with institutional and departmental leadership was and is chosen by EVA as a strategy for organising the individual evaluations more in detail. This process is "real" in the sense that EVA often adjusts the organisation and design of a given evaluation after this round of consultation.

The system that was established in 1992 had the following procedures, which basically have been kept unchanged ever since (Thune 2001: 7-8).

- EVA conducts a *preliminary study* with the purpose of identifying relevant study programs to be included in a given evaluation (due to difficulties in knowing the content and profile of some study programs), and to establish possible criteria/objectives to be used as a mandate for the evaluation. The final selection of study programs in a given assessment is decided after the preliminary study.

Based on the preliminary study and internal guidelines, EVA prepares the terms of reference. The terms of reference is a formal basis for an evaluation and they have to be approved by the Board before the evaluation process can be started. The terms of reference include the background and purpose of the evaluation, time schedule, the list of higher education unites involved in the evaluation, items to be included in the evaluation, the division of responsibilities between the evaluation group and EVA, the general framework for the evaluation, and the methods to be applied (EVA 2002).

Traditionally, a fitness-for-purpose approach has been used in Denmark, emphasising the objectives of a given study program (EVA 2002: 14). This approach has been balanced by including national policy objectives in the mandate when relevant. Hence, each evaluation conducted has a specific mandate. In previously conducted programme evaluations the following items were included:

- the objectives of the programme,
- management, organisation and resources,
- structure of the programme,
- content of the programme,
- practical learning,
- methods of teaching and training,
- lecturers/professors, including pedagogical competencies,
- exams and evaluation of students,
- students entry levels and progression,
- internationalisation,
- relations to other institutions and society,
- quality assurance.

The physical environment and learning resources have also been regularly included. (EVA 2002:14)

When an institution is selected for a review, it is informed of the evaluation process in writing. It is mandatory that EVA informs the institution on the legal basis for the evaluation, including the rights and obligations of the institution, the purpose of the evaluation, the terms of reference, the

members of the evaluation group and expectations to the institution's own contribution to the process (EVA 2002).

- The unit responsible for offering a given study program (usually a department) then writes a *self-evaluation report* based on a rather detailed protocol provided by EVA. This protocol usually instructs the department to describe the objectives of the program, the management and organisation surrounding the program, content, methods of teaching and learning, and to include quantitative information concerning the number of academic staff, applications, drop-outs, completion rates, etc. for the last three years. The purpose of the self-evaluation is both to provide the external expert committee with background information and to stimulate 'development' in the department. The department is free to choose the organisation of the self-evaluation process, but is advised to include academic staff, administration and students in the process.
- In parallel to the self-evaluation, a comprehensive *survey* on the quality of the programs is often conducted among various users, i.e. students, graduates, and employers (and sometimes external examiners within the subject field). These surveys are outsourced by EVA to private consultancies, poll firms, etc. The purpose of the survey is to provide alternative views and perspectives on the subject field and on each particular study program, and to indicate the relevance of the study program to important stakeholders. Due to the fact that these surveys are somewhat expensive, only one group of users are focused in each evaluation. The selection is made by EVA. Due to the fact that these surveys are conducted in parallel with the self-evaluation, a given department cannot use the results in their self-evaluation. The timing of the process is usually such that the self-evaluation and the surveys are finished simultaneously for the use of the external expert committee.
- The next phase in the evaluation process is the *visit* to the department *by the external expert committee*. Before the visit, the committee meets to discuss the information available, to find areas for investigation in addition to, or supplementing, the check-list used for the self-evaluation, and to plan the visit.

The committee is selected and chosen by EVA, and usually consists of three to five persons with extensive knowledge of the subject field and/or expertise in university governance and management. Committee members must be independent of the institutions involved in the evaluation, which is ensured by a formal statement from potential experts. The potential expert must state whether he/she has been employed, been invited to give lectures or in any other way associated with the programme under review, as well as whether his/her child, spouse or near friend has studied or been employed by the program (EVA 2002). Due to the fact that Denmark is a small country, where it is difficult to find "independent" experts, and that people from Denmark, Sweden and Norway can understand each other's language, experts are often recruited from the latter two countries. Even if these experts could be regarded as "peers", a committee usually also includes a member (sometimes two) from outside of higher education. Typically, this representative is from business, industry or a public organisation. The expert committee must be approved by the EVA's board.

A visit to a department usually takes place over a two-day period where representatives of the leadership, academic staff, administration and students are interviewed. Usually, but dependent on the number of study programs evaluated, the same committee undertakes all the visits. After the site visit, the committee meets two or three times in order to discuss and finalise the evaluation report.

For each evaluation, EVA appoints an internal project team of two evaluation officers and one assistant, which provides secretarial help to the external committee and ensures that the evaluation is conducted as specified in both the terms of reference and EVA's formal regulation. The project team is involved throughout the process: from conducting the preliminary study, and preparing the program and interview guides for the site visit, up to drafting the report to be discussed and approved by the evaluation group (EVA 2002).

- After the visit, and after several meetings in the external expert committee, an *external report* is written presenting an overall analysis of the quality of the program field at the national level as well as individual analyses of all study programs included in the evaluation. A draft version of the report is then sent to all departments/institutions involved in a given evaluation, and a closed conference is held where only representatives of EVA, the external expert committee and the departments/institutions participate. The purpose of the conference is both to prepare the departments for the coming conclusions, but also to provide an opportunity to adjust the report for any misperceptions or errors. After the conference, the final report is printed and sent to every participating department/institution, the Ministry of Education, and the relevant NEC. The self-evaluation report, the survey results and the external report are open and accessible to the public. The NEC has the responsibility to follow-up the report, for example, by checking how departments/institutions took actions based on the recommendations given in the report. It is important to note that the departments/institutions are not mandated to follow the recommendations, but will be asked questions related to what actions have been taken on the basis of the report.

The typical evaluation conducted by EVA includes up to ten study programs (occasionally even more) within a given subject field, with a self-evaluation report made for each study program/department, and with an external visit to every study program. After the release of the final report, Danish newspapers have a tradition for taking an interest in the results, sometimes creating a sparkling public debate.

Implementation

The implementation process in the first years was not without some turbulence. When the study program evaluations and EVA were proposed by the government, the Rectors conference, the umbrella body of Danish universities, was invited to participate in the development and design of the new centre. The Rectors conference, however, turned down the invitation from the Ministry. The argument used was that the Rectors conference preferred external evaluations, if necessary, to be carried out by the Ministry itself, and not by an independent body (Evalueringsscenteret 1998: 16). This scepticism from the universities' side can be related to both fear of loosening up established ties between the Ministry and the institutions (the established power structure), but also to protest since the institutions had to cover part of the expenses related to the evaluations themselves (the self-evaluation).

Due to the initial scepticism towards what the evaluations would bring, and the consequences of the evaluations, the procedures surrounding the evaluations were delicately designed (Askling et al 1998: 11). Because of the fear of being seen as just a cover for ministerial and political agendas, or to be perceived as "soft" towards the institutions, the procedures surrounding the evaluations had a focus on methods, systematisation and standardisation (almost unprecedented for quality assurance agencies) as the way to gain legitimacy and respect (Askling et al 1998: 12). Using EVA's own words: "the method is developed with the aim of uncovering the quality of a study program through a concrete, transparent and trustworthy process" (Evalueringsscenteret 1998: 25, authors translation). The fact that the mandate for any given evaluation is developed after a pre-study of the subject field, that rather detailed instructions are provided for the self-evaluation process (the protocol), that data is collected from current and graduated students and employers, and the arrangement of a conference with the involved parties before publication checking for any errors or potential overlooked problems, are all indications that the evaluation process is designed to be as robust and solid as possible. Since

private consultancy/poll firms were hired to collect and analyse the data from graduated students and employers, not only “methodological” triangulation was obtained (the use of different methods to shed light on a phenomenon), but also “observer” triangulation (the use of different actors to observe the phenomenon). Not surprisingly, all this information sometimes created very lengthy evaluation reports (up to 200 pages, see Thune 2001: 8).

Over the years, the evaluation procedure has been incrementally changed and developed. Some of the elements of the evaluation procedure have also evoked debate which sometimes has triggered adjustments. Issues that have been raised are, first, related to the protocol for the self-evaluation. In the first years, this protocol was very detailed with the potential effect that it was perceived as less relevant for initiating more developmental processes at the department level (Askling et al 1998: 16). Thus, over the years the protocol has been revised with the purpose of providing a broader framework for the self-evaluation instead of being a “questionnaire” to be answered. More open, reflective questions to be answered have been included in the protocol (especially related to how departments have established routines and systems for quality assurance).

Second, a debate has focused on the ability of a given evaluation to relate to the needs of the individual study program when the evaluation covers up to and sometimes even more than ten programs at the time. Stensaker (1999: 259) has, for example, documented how the departmental/institutional perceived benefit of an evaluation drops when more than ten study programs are evaluated at the same time. This study shows that the perceived benefit related to the self-evaluation process is constant independent of the number of study program participating in an evaluation, but the benefit related to the external panel visit and external report is perceived as smaller the more study programs are included. An explanation is probably related to a capacity problem in the external expert committee. A huge number of participating study programs results in more general recommendations from the committee, and the individual study program is not considered in the same way as when a given evaluation only covers a few study programs.

Third, the user surveys have also been criticised over the years. One line of criticism has been directed at the (lack of) competence of private consultancies for designing and analysing useful surveys (Evalueringscenteret 1998: 37). Another argument has been related to the timing of the user surveys in the overall evaluation process. Stensaker (1999: 263) has, for example, argued that due to the fact that self-evaluation and user surveys are conducted simultaneously, the departments cannot use the information from the surveys in their own self-evaluation. As such, it could be argued that these surveys have been more related to external (the external expert committee) than internal (departmental) needs. Since it is EVA in cooperation with the given consultancy/poll firm that decides the content of the survey, departments/institutions also miss a chance to put “their” issues on the agenda.

A fourth issue is related to the fact that a number of departments/institutions have been somewhat dissatisfied with how the closed conference works, and how the results of the evaluations have been commented upon in newspapers after the launching of the final report. The time for debate within the conference is limited and therefore the problems/potentials of each study program have not been fully addressed. In addition, many of the participating departments are also potential “competitors” in the Danish higher education market and have a more reserved attitude towards openly discussing problems and solutions (Evalueringscenteret 1998: 44, Askling et al 1998: 21). The fact that Danish newspapers, after the launching of the final report, tend to focus upon findings that give departments bad publicity, has also been raised as an issue, but the departments have not suggested shielding the external reports from public scrutiny.

Despite some debates and criticisms, the systematic evaluations were continued according to the plan. By 1999, EVA had fulfilled its mission and produced 62 evaluation reports, conducted way over a hundred user surveys of graduated students, employers, etc., and involved approx. 200 experts in the evaluation processes (Evalueringscenteret 1998: 30). The question to be asked is, of course, related to the impact of all this.

Impact

When the initial period was over, the Ministry of Education initiated an external evaluation of EVA. This process was conducted much in the same way as an ordinary study program evaluation. Hence, not only did EVA have to write a self-evaluation report describing and analysing the previous years, but also an external review of the methods, procedures and roles was conducted by experts from other Scandinavian countries (Askling et al 1998). A user-survey was also conducted, where a private consultancy firm asked representatives of the students, the higher education institutions (rectors, deans, etc.) and other stakeholders about their views on EVA's activities (PLS-Consult 1998). This section is based on this evaluation process, on a separate article where some of the data collected by PLS-Consult was re-analysed (Stensaker 1999), but also on observations from an independent study conducted by an American researcher in the field (Massy 1999).

Concerning methodology, trustworthiness and relevance of the study program assessments, Askling et al (1998: 4-6, 26) stated that the conducted evaluations had gained legitimacy, and that accountability and improvement actually were balanced, even if the improvement dimension could have been highlighted more. The rectors of a number of universities have reached similar conclusions (PLS-Consult 1998: 14). The high degree of systematisation, the stringent routines associated with each evaluation, and the various sources of data used to evaluate study programs were mentioned as important factors leading to this conclusion. The fact that the evaluation system created in Denmark did not integrate any performance indicator system into the evaluation is probably a factor leading to a positive attitude from the higher education sector.

The critique from Askling et al (1998: 25) was, therefore, more directed at what they saw was the weak point – the follow-up of the evaluations. Due to the “arms-length” steering principle, the Ministry only checked whether institutions had launched any actions after completion, and EVA had no responsibility for what happened after the publication of the evaluation report (Smeby & Stensaker 1999: 6). Thus, follow-up was a responsibility of the institutions themselves, but the external reinforcement for making sure actions were implemented was not great.

On this background, one could expect that the impact of the evaluations at the higher education institutions was limited. A survey of deans and department heads/other leaders at the department level, and a smaller number of interviews with rectors of various higher education institutions gives a rather different picture. Not least, a general attitude was that the evaluation processes had been a positive experience, and had created much discussion and dialogue inside the institutions (PLS-Consult 1998, Stensaker 1999, Massy 1999).

When asked to pinpoint the most beneficial element in the evaluation process, the self-evaluation process was undoubtedly most often mentioned (Stensaker 1999: 259-260). Many respondents also had a very positive view of the visit from the expert panel (PLS-Consult 1998: 15). The perceived benefits of the user-surveys were somewhat mixed, where feedback from students in general was seen as most beneficial (Stensaker 1999: 261). A negative comment mentioned by all respondents was that the external evaluations were time-consuming and rather expensive processes (PLS-Consult 1998: 14).

Concerning follow-up, slightly over 60 percent of department heads and other leaders at the department level claimed that recommendations had been followed up to a great extent, and that a huge majority of the respondents saw the recommendations given as good advices for improvement of the study programs (PLS-Consult 1998: 16). If one studies the areas where changes were most visible, the curriculum structure, examination, teaching methods, and the objectives of the study programs were the most often mentioned areas. Many respondents also claimed that the evaluations in general had triggered decisions and speeded up existing change processes in the evaluated departments (PLS-Consult 1998: 18).

In a broader, institutional, perspective, the assessments of study programs have had less effect. For example, for a number of years the protocol for departmental self-evaluation did not contain questions related to whether departmental systems for assuring and improving quality had been established. Hence, the assessments of the study programs contributed in this period little to the establishment of institutional routines for the systematic maintenance and development of the quality of teaching and learning (see, for example, PLS-Consult 1998: 17). Furthermore, since the protocol for the self-evaluations were fully developed by EVA, the departments participating in the evaluation often experienced less ownership, and hence less motivation for going into the evaluation process with an improvement orientation (Evalueringsscenteret 1998: 34). A last point to be made is that study program evaluations tend to de-couple the institutional leadership from the evaluation process. Even if the rectors often are drawn into the evaluation process, for example, by being interviewed by the external expert committee, one can detect a feeling that the institutional leadership perceive the study program evaluations as less relevant for them. For example, a majority of the rectors seemed to prefer an evaluation model where research and education were integrated in an evaluation (PLS-Consult 1998: 15). Another example is related to the fact that many rectors perceived that the costs associated with the evaluations outweighed the benefits (PLS-Consult 1998: 15). This perception changes, however, when those closer to the study program were asked about the perceived benefit. Almost 80 percent of the study program managers answered that they perceived the evaluation as relevant, providing them with valuable recommendations. In other words, greater distance to the evaluation creates less perceived benefit.

It is perhaps due to this de-coupling of the institutional leadership from the evaluations that Massy (1999: 30), analysing the national evaluation systems in Sweden and Denmark, stated that such systems “need not be an exercise in power and control”. (One should, however, also bear in mind that the traditional weak role of the institutional leadership in the Scandinavian countries could impact substantially on the conclusions drawn by Massy.) He maintains that one of the success criteria for these systems is that issues of power and control are not allowed to dominate the agenda. Perhaps as an effect of that, Massy (1998: 31) acknowledges that the national evaluation system has not transformed higher education. But, they have managed to “start universities and departments on the road to becoming learning organizations (...), to become self-conscious about the processes of teaching and learning and how to improve them”.

Costs

Costs have – on the national level - traditionally been a non-issue related to the evaluation system in Denmark. For example, the question of whether the country got “value-for-money” when a new system was designed in 1999 was totally ignored. The necessity to maintain a national system of evaluation and to expand it to all levels was dominating the agenda. However, as mentioned earlier, at the institutional/department level, complaints have been launched that participating in the study program evaluations is time consuming and economically burdensome since departments have to pay for the self-evaluation themselves. Finding resources for freeing staff to be included in the self-evaluation process has been perceived as somewhat difficult.

The costs related to funding and running EVA is decided by the Ministry of Education. EVA’s latest annual budget (2004) sums up to 43.6 mill. DKK (approx. 6 mill. US dollars). This includes the costs for evaluations in the primary school sector (both the primary and secondary level of education), but so far only a few pilot-projects have been launched in this area resulting in very limited expenditures.

The majority of the budget is covered by an appropriation from the Ministry of Education (approx. 40 mill. DKK). The rest of the budget, about 4 mill. DKK, stems from other external sources.

This budget covers the salary of 48 full-time employees at EVA, and ten to twelve student assistants working part-time. EVA estimates that the costs of a typical study program evaluation (dependent on the number of study programs included) vary between 250.000 and 500.000 DKK (approx. 33.000 – 66.000 US dollars). Most of these resources are tied up in travel expenses and honoraria for the external expert committee. A typical honorarium for an external expert, depending on the number of study programs included in the evaluation, varies between 12.000 and 25.000 DKK (1.400 – 2.800 US dollars). In addition, the costs of user surveys vary between 50.000 and 200.000 DKK (approx. 6.500 – 27.000 US dollars). As mentioned before, private consultancies are often the contractors and conductors of these surveys. In the first years, some evaluations became more expensive than budgeted and some were even substantially delayed according to their time schedule, but this problem has in later years been solved with more rigorous steering of each evaluation project (Evalueringsscenteret 1998: 24).

Comparisons, conclusions and future action

After the evaluation of EVA and the evaluation system in 1998, the Ministry of Education argued that systematic study program evaluation on the national level should continue, and that EVA had to be a permanent centre (UFK 1998: 33-36). However, the Ministry suggested that activities should be expanded, both concerning scope and practices. Hence, compared to the past, several new elements are visible in the current evaluation system in Denmark.

- Not only study programs in higher education, but education given at all levels in the Danish system are at present object for evaluation. The background for this expansion in scope was that Denmark experienced some (relatively) low scores in an international survey testing the knowledge level at certain ages in primary and secondary school (the PISA-survey), and that, amongst other things, more systematic evaluations were needed as a means to improve quality (Kristoffersen 2003: 26).
- Concerning methods, systematic evaluations of study programs in higher education are continued, but new forms of evaluations, including thematic evaluations, evaluations of institutions, system evaluations, and audits (evaluation of quality assurance systems) are also to be conducted (at all levels) (Kristoffersen 2003: 27).

The study program evaluations carried out over the previous years are still recognisable in the new version. This has probably a lot to do with the purpose of the evaluations, which still maintain the duality of accountability and improvement, and that the established procedures for conducting study program evaluations still are used. In a recent impact-study of some of the evaluations conducted by the new body, findings suggest that the study program evaluations continue to have an impact on Danish higher education, even in “round two” (Stensaker 2004: 38). The respondents do still have a positive view of the evaluations, approx. 60 percent of the recommendations are followed up, and a typical effect is that dialogue, discussions and reflections at the institutions increase afterwards (EVA 2004). The result is interesting in that counter to the predictions of researchers in the field, there are no signs of a diminishing impact of the national evaluation system as the institutions learn the “tricks of the trade” the second time around (Jeliazkova & Westerheijden 2001: 1). A possible explanation is that evaluations that stimulate dialogue, discussions and reflection may not be so vulnerable for that sort of institutional “strategic behaviour”. As neatly said by a US-observer: “The strength of the Scandinavian evaluation philosophy lies in its strategy of creating discourse rather than laying down regulations” (Massy 1999: 29).

However, the type of discourse that Scandinavian evaluations create, varies somewhat between the countries. In Sweden, where institutional audits (evaluations of institutional work to secure and improve quality) for a number of years were the dominant evaluation type on the national level, the discourse (measured in the number of recommendations made) has traditionally centred around management and strategy issues (Stensaker

2000). Those experiencing the effects of these audits perceived the problem that the audit did not address quality issues on the “shop floor”. Whether the quality of a given study program or department is good or bad, in other words, can not be detected by using this method. Contrary to this, the Danish evaluations address quality issues related to the teaching and learning process more directly. Most recommendations in the study program evaluations in Denmark have addressed issues related to teaching, curriculum and pedagogy, and organisation, management and strategy issues relating to quality have been less emphasised (Evalueringsscenteret 1998: 26).

That being said, it is important to remember that even if the focus of the Danish evaluation system is on study programs, it is still a “fitness-for-purpose” process and not directed at securing academic standards per se. The study programs in Denmark are not, as has been the case in the UK, checked against predefined academic criteria developed for every subject area, and neither have the same disciplinary orientation as the program assessments that have been carried out by the VSNU in the Netherlands (Brennan and Shah 2000: 64). A reason is probably the existence of the external examiner system in Denmark. In this way, one may argue that Denmark actually has two parallel systems for academic quality assurance on the national level. That these two systems in the past have been poorly linked has surprised external observers (see Dill 2002).

Dill (2002: 28) has also noted that the Danish study program evaluations have only to a limited extent focussed on assuring the quality of new study programs established in Danish higher education. As such, the Danish study program evaluations have been more oriented towards ‘general’ accountability in a system characterised by a relatively high degree of stability than targeted at quality assurance in a fast moving, more competitive and open higher education market. For example, the study program evaluations have been difficult to utilise as ‘consumer guidance’ for students. It can also be questioned whether they actually address political needs for information about higher education (Rasmussen 1997b: 257). The ‘accountability’ generated by the evaluations have, in other words, been directed more to ‘the environment’ rather than to a specific interest group.

On this background, one could ask whether the absence of “revolutionary change” in the Scandinavian higher education systems during the last ten to fifteen years (Finland excepted) may be one of the factors for the “success” of the Danish (and Scandinavian) evaluation approach. According to Massy, it is the “soft pedalling of the accountability agenda” that has provided the Scandinavian evaluations with high degree of legitimacy (Massy 1999: 33). Is it the absence of the need for radical political change of the higher education system that has provided a basic condition for the “soft pedalling”? Along the same line, it should also be noted that the Ministry of Education in Denmark has managed to follow up its own strategy and intention concerning the evaluation system, and that the legitimacy of the evaluation system probably also can be related to the fact that the Ministry have not interfered in the evaluation process. In other words, the arms-length strategy has been allowed to work according to the intentions.

Here, a contrast can be made to the Netherlands where the Inspectorate of Higher Education was to take action if a program experienced a poor assessment by the VSNU. This inspectorate, which reported directly to the Dutch Ministry of Education, could then instigate its own investigation potentially resulting in a deletion of the program from the list of recognised (funded) programs. Even if programs were not closed as a result of this procedure during the 1990s (Brennan and Shah 2000: 63-64), it is an indication of a policy design more oriented towards control than dialogue.

A comparison can also be made to the UK subject assessment system where the design of the subject assessments in the 1990s was more oriented towards grading different programs (accountability/control) than initiating improvement processes inside the institutions. According to Brennan and Shah, this relates directly to the “assessment methodology, in particular the summative judgement on a three-point scale” (Brennan & Shah 2000: 92). Typical effects of this methodology was the assignment and redistribution of status and reputation of programs, while a more typical effect of the Danish assessment methodology is that of changing institutional cultures and creating positive attitudes towards teaching and learning. Even if there were national funding schemes available in the UK for starting up quality development projects in teaching and learning at the institutions, the access to these resources was again dependent on a good assessment result (Brennan & Shah 2000: 94). Hence, to become better you should be rather good in the first place.

In retrospect, one may argue that the many policy problems highlighted in section two continue to exist in the new system in Denmark. Worries about efficiency, centralisation, control, development, relevance and how to deal with these problems are as present in Danish higher education today as ten years ago. The main difference is that it seems easier to articulate views on these issues today, and that the different actors trust more the motives and engagement of each other. The unfortunate accountability-improvement debate that in many countries, like the UK, dominated the policy agenda and consequently lowered the trust in the evaluation system was, after the first years, resolved through practice in Denmark (Askling et al 1998) and has not appeared since. Thus, Denmark is a good case for illustrating that external evaluation can balance accountability and improvement (see also Stensaker 2003: 157).

Seen in relation to the development in other countries, there has recently been an increased focus on international comparisons and on international developments within the field of quality assurance in Denmark. Accreditation is, for example, one of the new elements introduced in Europe in recent years (Faber & Huisman 2003), and EVA is currently working along a number of ways to “internationalise” their activities even further. Experimenting with audits of entire universities and colleges is one option currently pursued, yet another is to establish a procedure for mutual recognition of national quality assurance agencies (EVA 2002, Lindeberg & Kristoffersen 2002, Dill 2002). At present, there is in Europe a political drive towards more convergence of the systems of higher education (due to the Bologna-declaration), and consequently, more convergence of quality assurance systems as well. For the current Danish focus on study program evaluation, this seems to be the greatest challenge for the future. The problem is according to Thune (2001: 17), how a well-functioning (domestic) and established system of (“fitness-for-purpose”) study program evaluations can be translated or reinterpreted in the light of the (international) trend towards accreditation. How Denmark solves this problem remains to be seen.

Resources for policymakers

a) Live Internet Links

Information about the assessment of study programs in Denmark or the national evaluation system is available on the web-pages of the Danish Ministry of Education or the Danish Evaluation Institute:

- Ministry of Education: <http://www.uvm.dk/>
- Danish Evaluation Institute: <http://www.eva.dk/>
- The Danish educational system:
<http://www.eurydice.org/Eurybase/Application/frameset.asp?country=DK&language=EN>
- The official window to Denmark
http://www.denmark.dk/servlet/page?_pageid=85&_dad=portal30&_schema=PORTAL30&_fsiteid=175&_fid=12424

b) Technical documents

- Evaluation report in Agricultural Sciences (includes details about the methodology, style and process) ([pdf – 2,221KB](#))
- International Comparative Evaluation: Methodological experiences and future perspectives. ([pdf – 39KB](#))
- The EVA 2004 Action plan ([pdf – 302KB](#))
- Mutual Recognition: The Danish report ([pdf – 88KB](#))
- *Educational Evaluation Around the World. An international anthology.* See the chapter “Denmark” by D. Kristoffersen for a good overview of the current procedures, objectives and values related to EVA’s activities. ([pdf – 903 KB](#))
- Quality Assurance in the Nordic Higher Education: accreditation-like practices. ENQUA occasional paper 2

(pdf – 454KB)

- Follow-up on the Bologna-declaration – the establishment of a European Quality Assurance System (describes among other things the ENQA-network, a collaboration of European Quality Assurance Agencies) www.eva.dk -> publications
- The Ministry of Education action plan “Better Education” (pdf– 443KB)

c) Studies

Several of the studies in the reference list provide more information on the Danish evaluation system, see for example:

- Smeby & Stensaker (1999)
- Stensaker (1999)
- Massy (1999)

d) The new Danish evaluation system

The current white paper from 1999 (“Uddannelse og Fællesskab”) describing the background for and the elements in the new Danish evaluation system can be obtained from: <http://pub.uvm.dk/1999/uof/index.html> (in Danish)

References

Askling, B., Nordskov Nielsen, L., & Stensaker, B. (1998) *Mellom fag og politikk. En granskning av Evalueringscenteret og det danske evalueringsystemet for høyere utdanning*. Undervisningsministeriet, København.

Dill, D.D. (2002) Through the looking glass – Comments on the Nordic Pilot Project on Mutual Recognition. In Lindeberg, T. & Kristoffersen, D. (eds.) *A Method for Mutual Recognition. Experiences with a method for mutual recognition of quality assurance agencies*. ENQA Occasional Papers 4. Helsinki.

EVA (2002) *Mutual Recognition. The Danish report*. The Danish Evaluation Institute. Copenhagen.

EVA (2003) *Educational evaluation around the world. An international anthology*. The Danish Evaluation Institute. Copenhagen.

EVA (2004) *Effektundersøgelse. Redegørelse*. Danmarks evalueringsinstitut. København.

Evalueringscenteret (1998) *Selvevalueringsrapport*. København.

Faber, M. & Huisman, J. (2003) Same voyage, different routes? The course of the Netherlands and Denmark to a “European model” of quality assurance. *Quality in Higher Education*, 9(3), pp. 231-241.

Foss-Hansen, H. (1997) Universitetssektoren på vej mod år 2000. Den splittede omverden og behovet for helhed. I Rasmussen, P. (1997) *Universitetstradisjonen i krise*. I Rasmussen, P. & Jakobsen, A. (red.) *Universiteter i dag. Politikk – kultur – ledelse*. Samfundslitteratur. Fredriksberg.

Gornitzka, Å., Maassen, P. & Dahl Norgård, J. (2001) *Nasjonal prioritering og arbeidsdeling i høyere utdanning – internasjonale erfaringer*. NIFU skriftserie 9/2001. Oslo.

Gulldahl Rasmussen, J. (1997) Styring av forskning og undervisning på universiteter. Forandringer i de valgte mellomledernes funksjoner og plassering. I Rasmussen, P. (1997) *Universitetstradisjonen i krise*. I Rasmussen, P. & Jakobsen, A. (red.) *Universiteter i dag. Politikk – kultur – ledelse*. Samfundslitteratur. Fredriksberg.

Jeliazkova, M. & Westerheijden, D.F. (2001) *A next generation of quality assurance models. On phases, levels and circles in policy development*. Paper given at the CHER-conference in Dijon, 2-4 September.

Kristoffersen, D. (2003) Denmark. In *Educational Evaluation Around the World. An international anthology*. The Danish Evaluation Institute. Copenhagen.

Lindeberg, T. & Riemann, S. (2002) *Systematic evaluation in an international context: A small country perspective*. Paper presented at the 24th EAIR-Forum, Prague.

Lindeberg, T. & Kristoffersen, D. (eds.) (2002) *A Method for Mutual Recognition. Experiences with a method for mutual recognition of quality assurance agencies*. ENQA Occasional Papers 4. Helsinki.

Maassen, P. A. M. & Stensaker, B. (2003) Interpretations of self-regulations. The changing state-higher education relationship in Europe. In Begg, R. (ed) *The dialogue between higher education research and practice*. Kluwer academic publishers. Dordrecht.

Massy, W.F. (1999) *Energizing Quality Work: Higher Education Quality Evaluation in Sweden and Denmark*. National Center for postsecondary Improvement, Stanford University, Stanford.

PLS-Consult (1998) *Undersøkelse af effekterne af uddannelsesevalueringer*. København.

Rasmussen, P. (1997) Universitetstradisjonen i krise. I Rasmussen, P. & Jakobsen, A. (red.) *Universiteter i dag. Politikk – kultur – ledelse*. Samfundslitteratur. Fredriksberg.

Rasmussen, P. (1997b) A Danish approach to quality in higher education. In Brennan, J., de Vries, P., & Williams, R. (eds.) (1997) *Standards and quality in higher education*. Jessica Kingsley Publishers. London.

Smeby, J.C. & Stensaker, B. (1999) National quality assessments systems in the Nordic countries: developing a balance between external and internal needs? *Higher Education Policy*, 12(1), pp. 3-14.

Stensaker, B. (1999) User surveys in external assessments: problems and prospects. *Quality in Higher Education*, 5(3), pp. 255-264.

Stensaker, B. (2000) Quality as discourse: an analysis of external audit reports in Sweden 1995-1998. *Tertiary Education and Management*, 6, pp. 305-317.

Stensaker, B. (2003) Trance, transparency and transformation: the impact of external quality monitoring on higher education. *Quality in Higher Education*, 9(2), pp. 151-159.

Stensaker, B. (2004) Epilog: effekter af evaluering – en relativisering. In *Effektundersøgelse. Redegørelse*. Danmarks evalueringsinstitut. København.

Thune, C., Jacobsen, H. & Kristoffersen, D. (1996) Establishing an Evaluation System in Denmark. In Smeby, J. C. (ed.) *Evaluation of Higher Education in the Nordic Countries*. Nordic Council of Ministers. Nord 1996: 6, Copenhagen.

Thune, C. (1996) The alliance of accountability and improvement: the Danish experience. *Quality in Higher Education*. 2(1), pp. 21-32.

Thune, C. (2001) *Quality Assurance of Higher Education in Denmark*. Global perspectives on quality in higher education. In Dunkerly, D. & Wong, W.S. (eds.) Ashgate Publishers. Aldershot.

UFK (1998) *Kvalitet og ansvar. Fremtidens kvalitetsoppgaver inden for de videregående uddannelser*. Undervisningsministeriet, København.