



Evaluating higher education outcomes

Reflections on the Swedish evaluation system



Evaluating Higher Education Outcomes. Reflections on the Swedish evaluation system
Published by The Swedish Higher Education Authority 2014
Lars Haikola

The Swedish Higher Education Authority, Box 7703, 103 95 Stockholm
Tel: 08-563 085 00. Fax: 08-563 085 00
E-mail: registrator@uka.se, www.uka.se

Contents

Background	5
The Swedish evaluation model	6
The HEIs	6
The Swedish Higher Education Authority	6
A five-stage model	7
Characteristics.....	9
The evaluation system	10
Panels of assessors	10
Quality 10	
Selection of qualitative targets	10
The material evaluated.....	11
Methodology.....	12
Previous changes in the evaluation system	13
Quality aspects.....	13
Site visits	13
Questionnaires	14
Small programmes	14
Experience of the approach	14
Selection of targets	15
The size and complexity of the evaluation system.....	15
Recruiting assessors	15
Self-evaluations.....	16
Student interviews.....	16
Consultation meetings.....	16
Experience of the system	16
The critical debate.....	16
The evaluation system functions.....	17
The evaluation system is fair.....	18
Development	23
<i>Resultatsök</i> and information on which students can base their choices.....	27
Conclusion.....	28
Experiences unrelated to the system	28
The value of qualifications.....	28
Self-contained courses.....	29
Qualitative targets	30
How many programmes are covered by the evaluation system?	31
Main field of study	33
Double degrees.....	34
Calibration36	
Transparent but mechanical system	37
ENQA 38	
Resources	40
Looking ahead.....	41
Development idea 1: Broader results.....	41
Development idea 2: A greater wealth of information	42

Developmental idea 3: More scope for the experts	42
Development idea 4: Evaluate the HEIs' quality assurance system.....	43
Development idea: 5 More explicit labour market preparation	43
Development idea 6: Minimise homogenisation effects.....	44
Development idea 7: Augment calibration	44
Development idea 8: Enhance the student perspective	44
Concluding reflections	45

In 2011 Högskoleverket (the Swedish National Agency for Higher Education) launched a new evaluation system for higher education in Sweden. This task, which involves evaluating courses and programmes that lead to the award of first and second-cycle qualifications during the period 2011–2014, was taken over by Universitetskanslersämbetet (the Swedish Higher Education Authority - SHEA) in 2013.¹ Today therefore almost the whole cycle has been completed. This report presents an account of the experiences gained during the work devoted to the evaluation system over three and half years.

Background

The Bologna Declaration and the ensuing process of reforming higher education in Europe stresses the importance of organising education in terms of expected learning outcomes. Comparability, recognition of qualifications and quality assurance are to be based on shared targets rather than how educational programmes are organised. The expected learning outcomes are laid down in Sweden in a qualifications ordinance and are used in the way evaluation is conducted. The Government bill referred to as the ‘Bologna Bill’ stated the following:

In addition the link between the Qualifications Ordinance and the national quality assurance system should be explicit.

(Ny värld – ny högskola [New world – new higher education]. Government Bill 2004/05:162, p. 97)

The Bologna Bill also emphasised how important it was for the HEIs to be free to choose their own means of attaining the qualitative targets.

... the targets and other requirements laid down in the qualification descriptors can be attained in different ways and through different programme structures and the HEIs should have both the freedom to choose different strategies for attaining these targets and the responsibility for doing so. This principle should also apply in the future to all qualifications.

(Ny värld – ny högskola [New world – new higher education]. Government Bill 2004/05:162, p. 100)

The Government Bill submitted in 2009/10 entitled Fokus på kunskap – kvalitet i högre utbildning [Focus on knowledge – quality in higher education] (Government Bill 2009/10:39) placed emphasis on the importance of the development of different strategies for attaining the qualitative targets by the HEIs as the extent of their autonomy increased.

¹ Exception: Two-year general qualifications or two-year qualifications in the fine, applied or performing arts. Third-cycle programmes are not evaluated in this system.

Greater autonomy entails greater responsibility for the HEIs to structure their activities so that high quality is developed and sustained

(Fokus på kunskap – kvalitet i högre utbildning [Focus on knowledge – quality in higher education] Government Bill 2009/10:139, p. 7)

The point of departure for the SHEA's task is that the HEIs are responsible for developing their own activities, including responsibility for quality assurance.

The task of the Authority is to ensure that programmes and courses maintain high quality standards by means of a national quality assurance system that focuses on outcomes.

The Swedish evaluation model

The Swedish evaluation model is based on the assumption that quality procedures are in operation at both the HEIs and SHEA.

The HEIs

The HEIs are responsible for the quality and the development of the courses and programmes they offer. Most of the work on quality assurance procedures is undertaken by the HEIs. Their operations are regulated by the Higher Education Act, which stipulates that everything they do is to be of "high quality" and that quality procedures are to be undertaken jointly by their staff and students.

The operations of higher education institutions shall be arranged to ensure that high standards are attained in courses and study programmes as well as in research, and artistic research and development.

The resources available shall be used effectively to sustain a high standard of operation.

Quality assurance procedures are the shared concern of staff and students at higher education institutions.

(Higher Education Act Chapter 1: 4)

The Swedish Higher Education Authority

The role of SHEA in the Swedish evaluation system is to assure the quality of higher education. Its instructions stipulate that the Authority is to be responsible for the quality assurance of higher education through:

- Evaluations of first, second and third-cycle courses and programmes, and
- Appraisal of questions regarding degree-awarding powers pursuant to Sections 11, 12, 14 & 15 of Chapter 1 of the Higher Education Act (1992:1434) as well as Section 7 of the Act Concerning Authority to Award Certain Qualifications (1993:792).

(Swedish Higher Education Authority's Instructions, Section 2)

The aims of SHEA's evaluation model can be summarised as

- controlling that all higher education maintains high quality standards
- contributing to the development of higher education in Sweden
- providing information on which students can base their choice of programmes.

The first two of these can be found in all evaluation systems and these two objectives, quality assurance – development, and striking a balance between them is a constant challenge. In international discussions the balance is expressed in terms of accountability – enhancement. In the Swedish evaluation system there is emphasis on quality assurance/accountability. The quality evaluations undertaken by SHEA are at the same time expected to support development of higher education. An important element in today's higher education community involves enabling students to make rational choices among offerings in which the quality of the courses and programmes is a vital factor. The third aim of the evaluation model is to provide students with information on which they can base their choices. These three aims relate on the whole to three different target groups – the Government which allots the funding, the HEIs and the students. In addition to these three aims it can be added that the evaluation model must also meet needs expressed by the labour market and the social partners.

A five-stage model

The Swedish evaluation model comprises five stages.

1. Development work and applications
A HEI plans a new programme and therefore undertakes extensive development and quality assurance work. It then applies for entitlement to award the relevant qualification.
2. Appraisal of degree-awarding powers
SHEA appraises the HEI's application. A peer-review group analyses the qualifications of the teachers, learning outcomes and programme syllabus, quality assurance procedures, etc. The group recommends a decision to be made by the University Chancellor.
3. Programme and qualification
If the University Chancellor decides to grant entitlement to award the qualification, the HEI offers the programme and awards qualifications to its students. At the same time the HEI continues its work on quality assurance procedures.
4. Evaluation
SHEA conducts evaluations based on peer review of all first and second-cycle programmes of at least three years in length. The University Chancellor decides on the evaluation to be awarded using a three-level scale.
5. Follow-up
In cases where quality shortcomings are considered to exist a HEI is given the possibility of improving its programme within one year and accounting for its

measures. SHEA then undertakes another peer-review based evaluation of these improvements and makes a new decision.

Comment on stage 1. The Higher Education Act lays down (Chapter 1 Section 12) that a HEI that does not have university status must have been granted entitlement to award a third-cycle qualification and that a HEI must have been granted entitlement to award first and second-cycle qualifications. *This latter stipulation applies to all HEIs, including those with university status.* The act therefore stipulates that all HEIs must apply for entitlement to award all first and second-cycle qualifications – both general and professional. In practice this works differently.

When the 1993 reform of higher education was implemented the universities were allowed to continue to enjoy the degree-awarding powers to which they were entitled in the previous system. All degree-awarding powers for general and professional qualifications continued automatically to apply. This decision is still in force but is reviewed through the evaluation system. Moreover, in 2001 the Swedish National Agency for Higher Education decided on general entitlement to award 60-credit Master's degrees for certain HEIs that did not have university status. The model therefore means that in practice today all HEIs that do not have university status apply for entitlement to award mainly 120-credit Master's degrees and PhDs, and that all HEIs apply for entitlement to award professional qualifications.

Developing a new degree programme intended to lead to the award of a qualification is an extensive process that requires several years. This applies to all HEIs irrespective of whether an application for entitlement to award the qualification is submitted or not. Applications for entitlement to award qualifications presuppose that advanced quality assurance procedures are in place at each HEI.

Comment in stage 2. Appraisal of degree-awarding powers has been undertaken in the same way by the National Agency for Higher Education and SHEA since management by objectives in the form of the Qualification Ordinance was introduced in connection with the 'Autonomy reform' in 1993. (Universitet & högskolor – frihet för kvalitet [Higher Education Institutions – autonomy for quality] Government Bill 1992/93:1). Appraisal comprises evaluation of

- the circumstances in which the programme will be offered (teachers' qualifications, educational setting, infrastructure)
- implementation of the programme (policy documents, teaching, reading lists and examination)
- programme outcomes (ensuring that qualitative targets are attained, quality assurance of the programme)

For appraisals relating to 120-credit Master's degrees, PhDs and qualifications in the fine, applied and performing arts these main headings are adapted. Appraisal of entitlement to award a qualification means that SHEA assesses whether the circumstances that prevail at

the HEI enable it to offer a programme that has a high standard. This entails traditional accreditation of the programme and complies with an international template.

SHEA uses a peer-review based evaluation for its appraisals of entitlement to award qualifications. It should be noted in particular that one of the factors assessed is the standard of the quality assurance procedures that each HEI has developed and adopted. Appraisal of entitlement to award a qualification always includes assessment of the quality assurance procedures and system in place, irrespective of the qualification involved.

Comment on stage 3. Implementation of a programme is the most obvious stage and requires least explanation. A programme ends with its students applying for and being awarded a qualification. To qualify for this award all of the qualitative targets in the Higher Education Ordinance (Annex 2, Qualifications Ordinance)² must have been attained. HEIs may have different ways of ensuring that the qualitative targets have been attained but they are all based on sound quality assurance systems.

Comment on stage 4. This stage consists of today's evaluation system [The term "evaluation system" will be used to refer to stage 4 alone, while the term "evaluation model" will be used to refer to all five stages together.] This system involves *monitoring* an entitlement to award a qualification that has already been granted. One explicit aim of the evaluation system is that it must also provide a basis for the review of all degree-awarding powers. The main question on which the evaluation system can concentrate is: *How do you (the HEI) know that the students you have awarded qualifications to really attain the qualitative targets?* The question can be reworded in the imperative: *Demonstrate or make it likely that the students you have awarded qualifications to really attained the qualitative targets!*

Comment on stage 5: From 2013 onwards a fifth stage will be added to the model – follow-up. During the spring of 2012 the new evaluation model's first decisions were made and the quality of some programmes was judged to be inadequate. The HEIs affected were given one year for development work before they had to report on the improvements they had made. A new evaluation is then carried out, again on the basis of a peer review. Some form of follow-up has existed in all previous evaluation systems.

Characteristics

The Swedish evaluation model has two characteristics. The first is that it lays stress on the central and explicit role of the qualitative targets. Throughout the international educational world a shift is taking place towards the evaluation of results and emphasis on learning outcomes, which is in its turn an adaptation to the Bologna process. Sweden has uniquely positive conditions for undertaking outcome-focused evaluation as national targets have been laid down for all the qualifications in its Qualifications Ordinance.

² Separate ordinances govern the operations of Sveriges Lantbruksuniversitet [Swedish University of Agricultural Science) and Försvarshögskolan (Swedish Defence College). All subsequent references to the Qualification Ordinance include the corresponding ordinances that apply to these two institutions.

Secondly, the model emphasises that the bulk of the work on quality assurance procedures takes place at the HEIs. Considerable input is provided by SHEA but this is not as significant or as extensive as the internal work on quality assurance at the HEIs themselves.

Evaluation on the basis of qualitative targets is straightforward ideologically – it is reasonable to evaluate operations subject to management by objectives in terms of the targets that apply. But methodologically it is difficult to carry out this kind of evaluation. The number of targets for the qualifications laid down in the Qualifications Ordinance vary from 8 (Bachelor's degrees) to 24 (Master's degrees in upper-secondary education). There are targets that are vague, ambiguous, overlapping or difficult to operationalise and measure. Experience shows unambiguously, however, that it is possible to carry out evaluations of higher education using this system.

The evaluation system³

Panels of assessors

SHEA uses external assessors for all forms of evaluation. Most of these assessors are experts within the subject area to be evaluated, in other words the method is a traditional peer review. In addition to the subject expert panels also include students and labour market representatives. Altogether the panels of assessors should have sufficiently extensive and in-depth expertise to be able to evaluate all the targets for the programmes involved. The aim is also to have an even gender balance and significant representation of experts from other countries. One restriction here is that the assessors must be able to read the independent projects – degree projects – and other material in Swedish. The HEIs are offered an opportunity to nominate subject experts. Labour market representatives are nominated in the same way. Student unions are notified of the impending evaluation projects and then students themselves indicate their interest in participating as assessors. The decision to appoint the members of the panels is made by SHEA.

Quality

SHEA bases its evaluations on an operational definition of quality. This is defined as the *degree of target attainment*. This means that SHEA reviews the extent to which the students' *actual* attainments correspond to the *expected* learning outcomes.

Selection of qualitative targets

The targets in the qualification descriptors are grouped under three headings or types of knowledge that are the same for all the programmes:

- *Knowledge and understanding*
- *Competence and skills*

³ How evaluation is carried out is described at <http://english.uk-ambetet.se/qualityassurance/standardevaluationprocess.4.4149f55713bbd917563800010219.html>

- *Judgement and approach*

In each evaluation project a selection of the targets is made for the relevant programmes. This selection takes into account the requirements for higher education laid down in the Higher Education Act (Chapter 1, Sections 2 & 8–9). In making the selection it is also ensured that the targets chosen will together reflect the programmes in their totality as well as their distinctive features and the employability of their students. The discussion of these selections begins in the panels of assessors, which then submit proposals to the HEIs concerned so that they can express their opinions. At the same time the panels of assessors discuss what criteria the targets should be subdivided into. These are also circulated for discussion with the HEIs concerned. Only then does SHEA determine which targets and criteria the evaluation will be based on. These are published on the Internet. There is a significant degree of peer-to-peer input in the process of selecting the targets and criteria on which an evaluation will be based.

The material evaluated

Three different types of material are used to evaluate whether the selected targets in the Qualifications Ordinance have been attained.

- the students' independent projects
- the HEIs' self-evaluations
- students' experiences

Initially a fourth form of material was included – alumni questionnaires – but these had to be eliminated, however, because of the usual problems with questionnaires (difficulties in tracing alumni's addresses after several years, low response rates).

The assessors must use these three types of material to look for symptoms that will convince them that the qualitative targets are attained.

The students' independent projects

The independent projects – degree projects – are an important element. A random selection is made of a maximum of 24 projects (marked and passed) for each of the programmes evaluated or at least 5. If fewer than five degree projects have been produced within a programme during a three-year period, this evaluation method is not used. The selected projects are anonymised so that no information about the author, the HEI or the grade awarded can be found. This appraisal by the assessors does not involve any review of the grade awarded by the original examiner.

The HEIs' self-evaluations

The HEIs' self-evaluations are intended to serve two purposes. The first, and most important, is to enable a broader and more comprehensive presentation of outcomes than that provided by the independent projects. This is necessary as the independent projects cannot cover all of the targets that have to be attained for a specific qualification. In their self-evaluations the HEIs therefore have to account for, analyse and evaluate the

outcomes attained for all of the targets the evaluation is based on. Some presentation of the circumstances in which the programmes are offered and their processes can be required to enable the HEIs to provide a reasonable account of how they ensure that the students actually attain these targets.

The second purpose is to account for and explain circumstances that have a manifest bearing on the outcomes of the programmes. One circumstance consists of the teaching resources used in the programme evaluated. HEIs can also describe special circumstances relating to the students, if they believe these affect outcomes.

Students' experiences

The students' experiences of the programmes they take part in are an important supplement to the independent projects and the self-evaluations. For this reason information is collected to illustrate the opinions of the students on the outcomes of the programmes evaluated in relation to the targets laid down in their qualification descriptors. This is carried out in the form of structured group interviews with students nearing the end of their programmes.

Site visit

The panels of assessors also conduct site visits. Nowadays these take the form of web-based interviews with representatives of the relevant programmes at the HEIs.

Methodology

The members of the panel with appropriate academic qualifications read the independent projects and assess whether they offer evidence that the selected targets have been achieved. Altogether this results in a matrix of 4 – 12 targets which is applied to 24 independent projects. For each target an assessment is made as to whether the independent project demonstrates very high, high or inadequate attainment of the target. Target attainment is also reviewed in the self-evaluations, student interviews and, at times, questionnaires. Impressions gained during the site visit also play a role. An overall assessment is then made for each target. Failure by individual independent projects to demonstrate attainment of several targets does not affect the final assessment. On the other hand, if one and the same target is considered not to have been attained in several independent projects, the overall assessment will be that the target has not been attained in the programme as a whole so that the overall evaluation for the programme may be “inadequate quality”. Shortcomings in one single target may therefore be decisive. If attainment of all the targets is considered to be high, this is a strong argument for deciding that the overall evaluation will be “high quality”. “Very high quality” can be the evaluation given if at least half of the targets together demonstrate very high target attainment.

The panel of assessors presents a final evaluation for each of the programmes evaluated using the scale ‘very high quality’, ‘high quality’ or ‘inadequate quality’. The evaluation also states how well the students have attained the targets evaluated. A brief explanation is provided by the panel for each of its evaluations. The report from the assessors constitutes a recommendation of the decision to be made by the University Chancellor.

SHEA is responsible for the overall calibration of evaluations to attain parity between evaluation projects and the subject areas.

Evaluating a programme as having “inadequate quality” means that extension of its entitlement to award the qualification concerned is only conditional. The HEI is given one year in which to improve the programme and to account for its measures, after which follow-up takes place. SHEA then carries out a new evaluation and either decides that the improvements are satisfactory so that entitlement to award the qualification is extended or the measures are considered inadequate. After it has been given a chance to express its standpoint, the HEI may have its entitlement to award the qualification withdrawn.

Previous changes in the evaluation system

During development of the system some changes have been made. Four of them are described below.

Quality aspects

The Government Bill (Bill 2009/10:139) and the Government decision (2010-07-08, U2010/4164,3546,3552/UH) governing this process stipulated that “the National Agency for Higher Education shall evaluate to what extent *actual* learning outcomes are attained in relation to the *expected* learning outcomes laid down in the Higher Education Act and the qualification descriptors”. The expected learning outcomes are defined as *quality aspects*. These quality aspects were therefore to be given their own wording in addition to the wordings in the Higher Education Act and the wordings in the Qualifications Ordinance. After an exchange of opinions with the HEIs, the National Agency decided not to produce its own formulation of the targets laid down in the Higher Education Act or Higher Education Ordinance. Instead the entire evaluation process has been implemented as one evaluation using the targets laid down in the Qualifications Ordinance. This was a well-considered change. Evaluating in terms of the targets that actually applied and not in terms of target descriptors whose status could be challenged endowed the system with stability. But at the same time one weakness encountered is that the qualitative targets are vague and ambiguous, so that subdividing them into criteria becomes important.

Site visits

The same Government Bill states that site visits are to be made. Site visits are an established method of gathering information during evaluations and always provide important knowledge and understanding of the distinctive features of programmes. During the very first evaluation projects physical site visits were made, some of them in the form of interviews with representatives of the HEIs in the National Agency’s/Swedish Higher Education Authority’s premises. At the same time a platform was developed for web-based meetings. This method was used with increasing frequency and soon took over completely. The main reason for the evolution of the site visits into virtual meetings was, of course, financial. Virtual site visits offer very significant savings – not only in terms of lower travel costs but also in reducing the time assessors have to devote to the evaluations. One precondition was that it had to be shown that one web meeting was

sufficient for important questions to be answered. It should be observed that this evaluation system follows a defined pattern and that there is no longer the need for as comprehensive a context as in the previous evaluation system. The questions that a panel of assessors wants to pose can usually be answered completely in a web meeting. Physical site visits still have a great deal to offer but their added value cannot be defended in financial terms. The representatives of the HEIs also declared that they were satisfied with web meetings as they too avoided the extra work of planning for a site visit with a series of meetings at different levels.

Questionnaires

One form of input listed in the directives for the evaluation system consisted of questionnaires to existing students and alumni. During the development phase it already became clear that questionnaires to existing students could not be expected to have a satisfactory response rate. After discussions with the student organisations, above all with SFS (the Swedish National Union of Students), it was decided not to embark on this aspect. It was replaced to some extent by the interviews with students in which their views on the outcomes of their programmes could be established. Attempts were made, on the other hand, in the first evaluation projects to send questionnaires to alumni. It turned out, however, that here too response rates were low and uneven. Problems arose, for instance, in locating students and sending them questionnaires several years after they had graduated. Both forms of student questionnaires had been phased out of the system by round four.

Small programmes

The evaluation system was planned to be used for programmes that lead to the award of a first or second-cycle qualification. At a very early stage discussion arose about how *small* a programme can be if the system is to continue to be reliable, i.e. how many degree projects were needed in an evaluation project for the assessors to be able to form an impression of the quality of the programme and how many were needed to avoid any one single assessor being given too much influence. The assessors expressed their scepticism about being able to make reliable judgements about a programme for which only a handful of independent projects were available and where they would have to turn to the other material, above all the self-evaluations, which could not be ascribed the same status. During the spring of 2011 the National Agency decided that the system would only be applied if there were at least five independent projects from the three preceding years. This threshold requirement excluded more programmes than anticipated (see the section *How many programmes are covered by the evaluation system? P.31*).

Experience of the approach

SHEA has evaluated almost two thousand qualifications⁴ so there is a wealth of experience of how the system functions. Initially some experiences relating to the concrete implementation of the system will be dealt with.

⁴ The evaluations deal with qualifications. A large number of programmes may lead to the award of the same qualification. Students may acquire their credits in either free-standing courses or study programmes. Two thousand qualifications comprises many times that number of programmes.

Selection of targets

One of the first tasks for a panel of first tasks is to discuss and propose which targets in the Qualifications Ordinance are to be used for each individual evaluation. This selection has to be adapted to the evaluation they are undertaking and the main field of study. The panel of assessors also has to discuss and propose criteria for their evaluation. The selection of targets and the criteria proposed are then presented for discussion in the “kick-off” meeting with representatives of the programmes to be evaluated.

The selection of targets and choice of criteria could be expected to be difficult and also to give rise to discussion at the meeting with the HEIs’ representatives but this has not been the case. The selection of targets has not led to controversy and neither has the choice of criteria. There seems to be a great deal of peer consensus about which targets are most vital in the various main fields of study and programmes and also about the criteria that should apply for the evaluations.

The size and complexity of the evaluation system

One important element in the evaluation system is its use of independent projects. This means that a great many independent projects have to be selected and processed using a special web-based application. In evaluations of major areas of study, such as business administration, sociology and professional programmes in nursing and engineering, many assessors are required to read large numbers of independent projects and account for target attainment in several dimensions. This phase of the process takes time and requires extensive resources.

The rest of the material – self-evaluations, student interviews and any questionnaires that have to be documented and assessed – is also extensive. The system involves a very large number of documents that all form the basis of the overall evaluation. The extent and complexity is much greater than in the previous evaluation system while at the same time increased transparency and comparability is required.

Recruiting assessors

SHEA has long experience of working with peer review and therefore of recruiting experts. There is also extensive experience of recruiting students and some experience of recruiting representatives from the labour market for assessment and evaluation. The evaluation system now in force is, however, as described above, very extensive and the need of teachers from Swedish and Nordic HEI’s correspondingly large. Nevertheless, recruitment has given rise to few problems, above all with respect to Swedish assessors, although it has been somewhat more difficult where Nordic assessors are concerned. An academic career presupposes readiness to undertake various kinds of assignments as experts or referees and evaluation is considered to be the same kind of task. On the other hand it has been difficult to recruit students to be interviewed and labour market representatives as assessors in spite of the support provided by the student organisations and the social partners. One explanation of this difficulty is that the evaluation system does not primarily meet any of the needs of these groups in an evaluation of higher education. A more banal reason is certainly lack of time.

Self-evaluations

To begin with there was no experience of how a self-evaluation should be structured and this uncertainty caused problems for the HEIs. SHEA's guidelines on self-evaluation enabled the gradual determination of a praxis which gave self-evaluations a more concrete form. During the first hesitant round the self-evaluations were mainly used to clarify contexts but, as will be discussed in more detail below, they have later provided an important basis for the assessment of target attainment. This applies in particular to programmes leading to the award of professional qualifications.

Student interviews

The student interviews have varied in character and this makes it difficult to make comparisons between programmes and evaluation projects. It has turned out to be difficult to recruit students who choose to devote time to interviews about their studies. The same is almost certainly true about the difficulty of recruiting students as assessors. Student interviews have occasionally had to be omitted completely. The students also find it difficult to assess attainment of qualitative targets that are sometimes unknown to them. Nevertheless, the student interviews have played an important role when they provide valuable input in preparation for meetings with representatives of the HEIs and have generated questions that can be posed to those responsible for the programmes.

Consultation meetings

In order to shed light on the relevance of the programmes for the labour market and the community it is necessary to involve representatives of the labour market in the evaluations. This is achieved both by having representatives of the labour market in the panels of assessors and also through the consultation meetings arranged by SHEA with representatives of the labour market before the evaluations begin. At these meetings there is discussion, for instance, of the qualitative targets for the programmes to be evaluated and the panels of assessors are later provided with summaries of what has been discussed.

Experience of the system

Now that the evaluation system has almost completed its four-year cycle important experience has been acquired. Some things could be foreseen but others have been surprising. One outcome of the evaluation system is that important aspects of higher education in Sweden have been illuminated in an unexpected way.

The critical debate

The Swedish evaluation system has been subjected to criticism. It is obvious that evaluation of an activity will be examined critically by those affected by it. The criticism can be categorised in one of three ways.

1. It is based on ignorance or misunderstanding of the system's actual function
2. It is correct and refers to something inherent in the system
3. It is correct, but what it refers to is not inherent in the system and can subsequently be remedied.

The intensity of the criticism cannot, however, be explained by saying that “critics must expect to be criticised themselves”. Instead two additional causes can be adduced to explain it.

1) Some of the criticism of what is inherent in the system illustrates that different ideologies prevail for evaluation. This applies above all to the criticism expressed by ENQA (the European Association for Quality Assurance in Higher Education) and which takes as its premise that all evaluation should take the targets and quality procedures of the object of the evaluation itself into account (see the section headed *ENQA* p. 38). The Swedish evaluation system does not take as its starting point – as ENQA demands – the targets and quality assurance procedures of each individual HEI but carries out the evaluation on the basis of external criteria, the Qualifications Ordinance. Similar criticism has been expressed in Sweden to claim that the inspection element is too strong in relation to the developmental features. The HEIs have reported difficulties in integrating the external evaluation system with their own quality assurance procedures. The overall effect is greater alienation of the Swedish evaluation system from the higher education sector and it is not viewed as one element in the developmental process in higher education. SHEA has become more of an observer and judge than a colleague, and assessment experienced more as a threat than an aid to development.

2) The Swedish evaluation system makes an impact. It makes a real difference both with regard to shortcomings and praise and at the same time its outcomes are explicit. Conditional extension of entitlement to award a qualification draws a great deal of negative attention to a HEI – in the mass media, among students and in the organisation itself. The positive attention aroused by being evaluated as attaining very high quality is not at all as great. The impact should be viewed against the background of the size of the sector. Over half a million individuals are involved in higher education in Sweden and it has a total turnover of SEK 74 billion per year. In future accountability for outcomes is likely to be greater than it is today – not less. Merely accounting for what system a HEI has to guarantee high quality will no longer do.

The ideological differences and the impact of the system and its effects can explain the harsh tone that is at times heard in the debate on the Swedish evaluation system. But the most important question to ask is whether the system serves the purposes it claims to. These were stated above as:

- controlling that all higher education maintains high quality standards
- contributing to the development of higher education
- providing information on which students can base their choice of programmes.

The evaluation system functions

We can establish that the system functions. This means that the system works logistically, as timetabled and operationally. As pointed out in the section entitled *The size and complexity of the evaluation system* (p. 15), the great extent of this evaluation system has to be emphasised. Almost 2,000 programmes/qualifications have been evaluated. Each programme has been evaluated in terms of from four to up to twelve qualitative targets. This means that many thousand individual targets have been assessed. A panel of assessors consisting of subject experts, students and labour market representatives,

totalling 5–20 members, has been recruited for each evaluation project. More than 1,200 external assessors have been appointed. Their material has comprised almost 2,000 self-evaluations, as many interview summaries and sometimes questionnaires. Altogether this comprises a very large amount of data that has been dealt with using a relatively stringent system in which timing has been important. In addition the various projects have been calibrated so that the assessment of totally different programmes evaluated over four years can be placed on the same shared scale. This very extensive and complex system has to all intents and purposes functioned. Only a few hitches and errors have arisen – usually in connection with an additional complicating factor, for instance when a programme leading to the award of a double degree has been included.

The question of calibrating all the projects is dealt with in the section entitled *Calibration* (p. 36).

The evaluation system is fair

What is more important is that the system provides a reliable measure of the quality of a programme. Measuring quality is admittedly difficult and can be carried out in several ways. The method adopted assumes that education is an activity that should lead to predetermined targets. These targets are laid down in the Higher Education Act, the Higher Education Ordinance and the Qualifications Ordinance. Here quality is defined operationally – the degree of attainment of the qualitative targets. Measuring quality by ascertaining to what extent the targets of an operation are attained is *one* reasonable method of determining the quality of a programme.

There are other ways of measuring the quality of a programme and evaluations of programmes are undertaken using other systems. For instance there are evaluations of quality made by Svenskt näringsliv (The Confederation of Swedish Enterprise) that measure quality on the basis of how long it takes for a graduate to enter the labour market and evaluations that assess quality on the basis of the salaries paid to graduates from different programmes and HEIs. One aspect that the Swedish National Union of Students wants to include in assessments of quality is the pre-entry qualifications of students in order to see what progress they make during their studies. There are organisations that want to evaluate the contribution made by programmes to equality of opportunity and gender equality. Some stakeholders assert the value of internationalisation and third stream activities and claim that the evaluation system should take these into account. There are those who advocate a return by SHEA to evaluating the HEIs' own quality assurance procedures. *All these alternative ways of determining the quality of programmes are interesting and reasonable.* They reflect *what* is evaluated on the basis of different stakeholder's needs. Measuring on the basis of programme targets is an expression of the need felt by the Riksdag and Government to establish clearly that the targets laid down and for which funds have been allocated are attained and to what degree. Undertaking several types of evaluation is fundamentally a resource issue. It would be best if all the stakeholders could divide the work up so that as many aspects as possible could be covered. The need of multifaceted and therefore extensive evaluations must be balanced, however, against the extra effort that would be demanded, especially from the HEIs.

Basing evaluations on the independent projects

It is difficult to find critics of the idea of measuring programmes in terms of the targets laid down for them. When this kind of criticism is expressed what is said is that the targets are not primarily intended for use in evaluations and not therefore appropriate (this is the reason why each target is subdivided into sub-goals and/or criteria) rather than that objective-oriented activities should not at all be evaluated in terms of their goals. The link between qualitative targets/qualification descriptors and evaluation is not new, however, but had already been expressed in the ‘Bologna Bill’.

Clearly worded qualification descriptors should offer the National Agency for Higher Education better conditions to make consistent evaluations of the quality of courses and programmes at different HEIs and also offer the Agency an appropriate instrument for the award and withdrawal of degree-awarding powers. Qualification descriptors should, therefore, be worded so that they can be used as far as possible for the quality assurance of courses and programmes.

(Ny värld – ny högskola [New world – new higher education]. Government Bill 2004/05:162, p. 99)

Greater criticism has been expressed of the methodology, i.e. how target attainment is measured. Here the main argument has been that the independent projects are given too great a role and it is not possible to assess the quality of a programme on the basis of the independent projects alone. The minimum number of higher education credits that can be awarded for the independent projects totals 15, which means that a three-year programme can end up being evaluated on the basis of the work of half a semester.

But the independent projects play a special role in a programme and cannot merely be viewed as one-twelfth of the whole. The Government Bill before the Bologna reform had already stated that the independent projects were to play a special role in reflecting an entire programme.

The Government considers that the independent projects are central in confirming that students have fulfilled the requirements for the award of a qualification. In their independent projects students demonstrate not only that they have acquired knowledge but also that they can apply and develop it with the degree of autonomy required to be able to pursue the career for which their studies have prepared them or to begin studies at a higher level. It is against this background that independent projects are being introduced for all first and second-cycle qualifications.

(Ny värld – ny högskola [New world – new higher education]. Government Bill 2004/05:162, p. 102)

When independent projects were introduced in all programmes in 2007 it was therefore linked to their capacity to confirm “that students have fulfilled the requirements for the award of a qualification”

Experienced teachers also know that there is a link between the other courses in a programme and the independent projects that usually conclude them. It is difficult if not to say impossible to complete a good independent project without having good results in the other courses that form the programme. Courses are structured with some degree of progression to lead towards the degree project, which is one of their functions.

In many programmes there is a long-standing and established tradition that students conclude their studies with a degree project, an independent project. This applies in particular in the humanities subjects that predominated in the first round of the evaluation. In engineering as well there is a tradition of degree projects that often, however, focus on practical industrial applications. In other programmes degree projects are also traditional. But for many programmes the independent projects were a new requirement from 2007. The programmes that have long experience of degree projects with theoretical contents have found it easiest to accept the evaluation system and have perhaps even benefited from it. Professional programmes have found it more difficult to reflect all their educational targets in an independent project. It is therefore representatives from these programmes that have expressed criticism of the predominant role played by the independent projects.

It was possible to predict this dilemma. Independent projects are often presented in written form (not, however, in the fine, applied and performing arts or architecture) and it is easy to realise that a written product cannot reflect all the targets of a professionally oriented programme. One example among many is that one of the expected outcomes laid down for Master's programmes in engineering is to "demonstrate the capacity for teamwork and collaboration with various constellations". There are also several qualifications that require *oral* skills (Bachelor's programmes in nursing, marine engineering, Master's programmes in law, etc. as well as general Bachelor's degrees) which cannot be demonstrated in a written text. For this reason the emphasis has gradually shifted from the independent projects towards the rest of the material, in particular the self-evaluations. One illuminating example of this can be found in programmes in dentistry, where the self-evaluations have been as important as the independent projects. Attainment of targets that it is difficult to substantiate in the independent projects must therefore be corroborated in some other way instead. This is one of the purposes behind the self-evaluations, as is made clear in the instructions about how they should be drawn up.

The self-evaluation comprises three sections. The first, and most important, is intended to enable a broader and more comprehensive presentation of outcomes than is provided by the independent projects.

(Report 2011:4 R Generell vägledning för självvärdering i Högskoleverkets system för kvalitetsutvärdering 2011 – 2014 [General guidelines for self-evaluations in the Swedish National Agency for Higher Education's quality evaluations 2011–2014], p. 5)

It has been suggested that perhaps the self-evaluation does not enable corroboration of attainment of a specific target either. A number of qualification descriptors contain the outcome "demonstrate the ability to identify the need for further knowledge and

undertake on-going development of his or her skills”. It is undoubtedly difficult to do so in an independent project or to corroborate that the goal has been attained in a self-evaluation as well. Here two points need to be emphasised. 1) When a HEI awards a qualification to a student it must reasonably know that the student has attained the qualitative targets – otherwise no qualification should be awarded. The evaluation asks about substantiation of the students’ attainment of the qualitative targets and for this to be accounted for in the self-evaluation. In the very nature of things, substantiation of aspects of this kind must be of a more processual nature and cannot be presented as isolated outcomes. 2) There may – in spite of everything – be cases when it is in fact impossible to demonstrate that targets have been attained. Then consideration should be paid to whether such a qualitative target should be retained in the future or whether we should quite simply accept that there are important objectives for higher education that cannot be verified or measured. These are important questions for the development of higher education and for its evaluation.

The criticism expressed concerning the role of the independent projects in evaluations can therefore be countered with the arguments 1) that there has been a requirement since 2007 that independent projects should reflect an entire programme; 2) that in most programmes this gives rise to few problems, above all in the traditional humanities programmes; and 3) that above all in professional programmes the role of the independent projects has de facto become less prominent and that other input has become more important in assessing whether qualitative targets have been attained.

The views of the assessors

The subject experts engaged for the evaluations have, above all in the first round which at the same time also involved a great deal of development work, simultaneously discussed the strengths and weaknesses of the system critically but have concluded that the results, evaluation on a three-level scale, give a fair picture of the programmes evaluated.

With reference to what is stated above about the evaluation system, targets and criteria, as well as the way it has been carried out, it is our impression that the appraisal, with the exception of the assessment of certain aspects of methodological and theoretical knowledge for the Bachelor’s degree mentioned, has led to a fair and consistent evaluation of attainment of the stipulated qualitative targets for the different qualifications. Our conclusion is that on the whole the evaluation has provided a sound and at the same time thought-provoking assessment of the condition of sociology and adjacent fields of study in Sweden.

Quality evaluation of sociology and adjacent fields of study. Panel of assessors’ report. Decision 29 May 2012

Similar summaries can be found in several of the panels of assessors’ reports. The view of the assessors is that primarily the system works well in differentiating between “inadequate quality” and “high quality”, in other words that the system is sound when it comes to ensuring that all programmes really do maintain high standards as required by

the Higher Education Act, but that it is weaker when it comes to the distinction between “high quality” and “very high quality”.

Early on in the process, the National Agency for Higher Education/ SHEA directed assessors to judge only target attainment to provide an overall judgement of a programme on this basis and to refrain from expressing opinions on how the system functions or other aspects of the programmes. It is therefore interesting that several of the panels of assessors have made spontaneous comments on the functionality of the system as well.

SHEA makes regular evaluations of the experts’ views on the evaluation process using questionnaires and these support the assertion that the system functions. The first questionnaire to assessors included the question “*How easy/difficult was it to assess the following qualitative targets with the help of the independent projects?*”. Where targets relating to knowledge and understanding were concerned, for instance “demonstrate knowledge and understanding in the main field of study”, 77.4 per cent of the assessors respond that it was very easy or fairly easy to assess this target on the basis of the independent projects. Self-evaluations were considered to provide a less reliable basis for the assessment of goal attainment. It should be pointed out that the first assessors’ questionnaire was based only on the experience of assessors in the first round. This comprised traditional “humanities” subjects where there is an established tradition of concluding with major degree projects. Undoubtedly this evaluation system suits programmes of this kind best and this was also the reason for choosing them for the maiden round of evaluations. Only one programme leading to the award of a professional qualification was included in round 1 – MSc in psychology. Professional programmes have successively formed a larger proportion of the following rounds. It is therefore interesting that in later assessors’ questionnaires confidence in the self-evaluations has risen markedly. In the first questionnaire 20 per cent of the respondents considered that the self-evaluation had played a very large role or large role for the overall evaluation. In later questionnaires this figure has varied between 40 and 60 per cent. The assessors in round 4 considered that the HEI interviews/site visits had played a significant or very significant role – the responses given by half of them. This reflects the need to assess professional programmes on a broader selection of material than the independent projects alone. In all the questionnaires more than 90 per cent of the respondents stated that the independent projects were very important or important for appraisal of target attainment. There is therefore still confidence in this element.

The assessor surveys are undertaken regularly and questionnaires have been sent to 1,200 assessors with a response rate of about 80 per cent. They provide unanimous corroboration that the system functions as a sound gauge of the quality of higher education.

Peer-to-peer assessment

International experts in the scientific council established as part of the evaluation system have pointed out that the Swedish system has a special strength in that it gives the assessors a role that they are expert in – judging whether qualitative targets have been attained in their own disciplines. Assessors are often selected because they are experts in their field, have sound teaching experience in the field and are good examiners. In many evaluation systems, including the one formerly used in Sweden, teachers are often asked

to assess systematic quality assurance systems at HEI level – something that few of them are experts on.

Reactions of the HEIs

One important indication that the system functions and is fair can be found in the reactions of the HEIs. When the evaluation system was introduced and the first results presented the system met with scepticism and criticism. In some cases the resistance to a critical evaluation was fierce and evaluations were not accepted. Reactions like this still occur today but they are not as prevalent. Gradually the validity of this new way of measuring quality has met with increased understanding, acceptance and acknowledgement. “Inadequate quality” is rarely an assessment that comes as a surprise. This reaction also embodies a pragmatic trait. Today the predominant reaction to an evaluation of “inadequate quality” is self awareness of the shortcomings of the programme in question. In some cases the criticised programme has even been closed down before the University Chancellor’s decision has been made. It is not uncommon for external evaluations to have a catalytic effect in accelerating a process that has already been initiated internally at the HEI. The new evaluation system and its results are also often viewed as important tools that enable administrators at various levels to adopt measures.

The reactions of the HEIs are also based on criticism of the lack of precision of the scale used. “Inadequate quality” is an assessment that tempts misuse in the mass media. It can be a severe and (far too) unambiguous way of summarising an evaluation that offers no scope for gradation. It is an assessment that fails to distinguish between a programme where there are shortcomings in all the targets and one which only fails to meet one of its targets – or even part of one. It is possible that the wording “questionable quality” would be more appropriate given the way the evaluation system is organised.

Development

In all evaluation systems a balance has to be struck between control and development and all systems comprise both aspects. The Swedish national quality assurance system is structured to place emphasis on control as the HEIs have the main responsibility for development. Despite this emphasis, the Swedish evaluation system includes explicit developmental elements, i.e. *the system is an incentive for quality*. Some of the developmental elements are deliberate and systematic, some of them follow in the wake of all evaluation irrespective of its design. A few of these effects have, nevertheless, been surprising. Among the first category are the developmental features referred to already in the section entitled *A five-stage model* (p. 7) in which the evaluation system constitutes a follow-up of applications for entitlement to award qualifications. In all activities that are subject to external evaluation, organisational measures are adopted that lead to enhancement of quality. The HEIs also testify unanimously that the self-evaluation process gives rise to insights and develops their operations. One of the surprising effects is, for instance, insight into the qualitative targets that the evaluation system has engendered. This is commented on specifically in the section entitled *Qualitative targets* (p. 30).

“Surplus information”

The question of surplus information should largely be viewed as an unintended effect. “Surplus information” means here that the evaluation system *contains* more information than it is required to *account for*. One of the premises of the system was that the task of SHEA was limited to reporting the extent to which qualitative targets were attained. In cases where inadequate quality is considered to exist, it is the responsibility of the HEIs to analyse why the programme has shortcomings and to remedy them. At an early stage there was critical discussion of whether this division of responsibilities could or should be strictly maintained. There are two reasons for reallocating responsibilities.

The first is that the panels of assessors obviously acquire considerably more knowledge about a programme that has been evaluated than the system provides scope to report. SHEA has explicitly urged assessors only to express opinions about whether the qualitative targets have been attained but, as experienced teachers, they have been able to get more out of the material than has been asked for. In some cases the panels of assessors have included this “surplus information” in their reports.

Secondly the HEIs have requested more information about their shortcomings. The accounts of the final reports presented under the heading *Resultatsök* on SHEA’s website have been considered far too skimpy to base any measures on. It was particularly problematic in round 1 to suggest remedial measures as no practice had been developed about what could be considered an adequate measure when follow-up took place.

This circumstance – that the panels of assessors have knowledge about programmes and those responsible at the HEIs have asked for additional information – must obviously justify striking a new balance between control and development. Consensus emerged in the sector early on that it should be possible to express the knowledge the panels of assessors had acquired and wanted to convey. In the current rounds the scope provided to supply “surplus information” has been increased but it still has no systematic role to play in the system. This should be integrated into the next generation of the evaluation system.

The panels of assessors possess not only knowledge about specific programmes but also form an image of the “state of the art” in the main field of study or a programme. Overall knowledge of this kind is exceedingly valuable for a subject’s or programme’s internal development. There is no system in Sweden to collect this “state of the art” awareness but many subjects and programmes are linked to disciplinary associations, professional organisations and the like that could assume responsibility for taking advantage of it. In some cases SHEA has initiated special conferences to enable the knowledge acquired by the panels of assessors to be exploited when there were otherwise no arenas for doing so.

At the same time it is important for the division of responsibility between SHEA and the HEIs to be upheld. The HEIs have the main responsibility for developing courses and programmes but SHEA’s evaluations could be used more effectively. How this balance should be attained is dealt with the section headed *Development Idea 2: A greater wealth of information* (p. 42).

Bench-marking

At an early stage of the evaluation cycle bench-marking effects could be observed. These could be internal as those responsible for programmes at the HEIs analysed evaluation reports and drew conclusions about what characterised programmes assessed as having inadequate quality, high quality or very high quality and then applied this knowledge to other programmes they were offering. This was true, in particular, of those responsible for programmes that were about to be evaluated in the following round. In other words evaluation of *one* programme at an HEI leads to developments in *several* more of the programmes it offers. Bench-marking has also taken place between the HEIs, above all where programmes assessed as having very high quality are concerned.

Follow-up

The final stage of the evaluation cycle consists of follow-up. Programmes rated as being of inadequate quality have one year in which they must report to the SHEA on the measures adopted to remedy the shortcomings revealed in the first evaluation. Follow-up applies only to the qualitative target or targets that were not, according to the evaluation, attained. One vital element is that it is the task of the HEI to analyse the shortcomings and decide on the appropriate measures required to remedy them. The report submitted by the HEI at the end of the year therefore consist of two parts – one its analysis of the shortcomings and the other an account of the measures adopted. The measures accounted for must be based on the analysis and already in place – not merely planned.

Peer review is used in this stage as well, i.e. a small group of experts (at least two and no more than five), often members of the panel of assessors that made the original evaluation, assess whether the measures described by the HEI are likely to remedy the shortcomings that prevented attainment of the qualitative targets, in other words it is future attainment of these targets that is considered. If the peer reviewers make a positive recommendation, the University Chancellor decides to award the assessment of high quality and the conditional extension of the power to award the qualification is made permanent. The assessment of very high quality cannot be considered when a follow-up decision is made. If the peer reviewers remain critical, their report is forwarded to the HEI for its responses. The University Chancellor can reach a decision on the basis of material from both the HEI and the peer reviewers and this can result in withdrawal of degree-awarding powers.

It is generally known that the evaluation system has resulted in just over 20 per cent of the programmes assessed being rated as having inadequate quality. Rounds 1 and 2, which started in 2011, led to decisions in 2012 and were followed up in 2013, included 581 qualifications and inadequate quality was the assessment made of 135 of them. Follow-up revealed that almost ten per cent of these programmes had been closed down by the HEIs themselves so that no remedial measures were accounted for. All of the other programmes were subsequently assessed as having high quality.

Criticism has been expressed by the HEIs of the way in which the follow-up evaluates something different from the outcomes on which the original evaluation was based. It has been shown above that it is impossible to restrict the material used to outcomes alone and that a significant shift towards the inclusion of processes and sometimes prerequisites has taken place, but nevertheless the fundamental idea is to evaluate outcomes – the

degree of attainment of qualitative targets. Follow-up cannot have the same focus on the outcomes of programmes as there can be no new outcomes after one year, and certainly not in the form of independent projects. Educational development normally needs more than one year to produce new outcomes. What the HEIs can account for are the measures adopted – which in the long run can lead to improved results. This means that different kinds of processes and sometimes prerequisites have to be assessed. This shift in the material on which evaluation is based may be viewed as problematic. But follow-up can also be seen as a supplement to the basic evaluation. The HEIs have asked for a reduction in the focus on outcomes and for greater weight to be given to processes – which is what the follow-up provides.

In the initial sections of their follow-up reports the HEIs, as described above, have to make their own analysis of *why* the qualitative targets were not attained on the basis of the SHEA's report. When the results of the evaluations are presented they contain no recommendations about suitable measures but the analysis of shortcomings and implementation of changes is left to the HEIs. In discussions of the system criticism has been expressed of the brevity of the SHEA's reports and the lack of feedback to those offering programmes. But the follow-up has revealed that, on the contrary, it has been manifestly easy for the HEI's to analyse the shortcomings for which programmes have been criticised. The self-analyses are on the whole insightful and often self-critical.

The next section of the follow-up report contains the HEI's account of the measures adopted to improve the programme. More than a hundred such measures have been listed and they are clearly linked to the analysis of shortcomings in each individual case. A number of measures focus specifically on the individual projects, such as introducing assessment templates for grading, more explicit grading criteria, clearer processes leading up to the individual project work, seminars for supervisors, seminars on methodology, cooperation between HEIs on individual projects and second-opinion systems for grading the projects.

But just as many measures focus on programmes in their entirety, such as new or totally revised syllabuses, new learning material, expanded reading lists, more intensive examinations, new courses, new course contents, compulsory literature searches, systematic contact with libraries, more explicit progression from course to course, revision that includes examination of all qualitative targets as well as clearer links to the surrounding community and ethical standpoints.

Some of the measures that directly affect students include more stringent entry requirements, an increase in the number of teaching hours, thorough changes of educational approaches, more stringent demands for individual projects and improved feedback to students. Many measures also involve enhancing contacts with research, linking individual projects to research projects and using researchers as supervisors. Other measures include employing more teachers, increasing the academic qualifications of examiners, limiting the number of individual projects one teacher can supervise, clarifying the division of responsibility between supervisors and examiners and matching students and teachers more effectively. Purely organisational measures such as establishing new boards or committees for programmes also occur.

One of the points for which the evaluation system has been criticised has been that there has been too much focus on outcomes that can be found in the individual projects, the HEIs' own self-evaluations and student interviews. Now that it is possible to survey the entire evaluation system, including assessment, improvement and follow-up, what can be observed is that the material used has been extended and that the balance between prerequisites, processes and outcomes has altered.

Some of the measures adopted by the HEIs gobble up resources which may have to be diverted from their other activities but this is far from true of all of them. Many involve clearer demarcation and operational improvements that can even lead to a reduction in the resources required.

On the whole the follow-up element in the evaluation system leads to a very large number of measures for improvement and a substantial increase in the quality of higher education in Sweden. This rise in quality is hardly likely to have occurred without the evaluation system.

Generally speaking, it can also be observed that the evaluation system has had systematic and powerful developmental features but some of its impact on development has not always been predictable.

***Resultatsök* and information on which students can base their choices**

The traditional conception of the factors that affect students' choices when it comes to their studies is that they are not directly linked to the quality of the programmes. They include how far away they are, the study environment, availability of accommodation, parental education, what friends are studying and where, as well as the programme's prestige. The major rise in the resources allocated to higher education in many countries has also led to greater demands for programmes to be implemented efficiently and to benefit individuals and the community. Now, therefore, students are increasingly expected to choose their programmes for different reasons. This expectation also seems to have had some effect judging from the survey of young people's attitude called *Ungdomsbarometern* (12/13). The latest survey reports that young people now select programmes *primarily on the basis of their quality*. This trend can be expected to grow in the future.

The results and documentation from all the evaluations are presented in a new form on a web application called *Resultatsök [Search results]*. This presentation has been praised in the debate. It provides a simple, clear and effective arena in which all of the evaluated programmes can be accounted for. It is easy to trace how an evaluation has been generated, who was involved, how the process was undertaken and what the evaluation was based on. This makes it simple to form an impression of a main field of study at a specific HEI or at all the HEIs and it also provides an overall view of all the programmes offered at a HEI. What is most important is that individual students can easily find information about the results of the evaluation of a specific programme at a specific HEI or at all the HEIs.

The third aim of the system was to enable students to base their choice of programmes on the results of the evaluations. *Resultatsök* is by far the most visited page on SHEA's

website. Even so it is too early to determine whether the third aim of the evaluation system has been attained through *Resultatsök*.

Conclusion

On the basis of this description of the extent of the material used (which comprises a great deal more than the independent projects alone), the views of the assessors and international evaluation experts and also, not least, the predominant reaction from the HEIs, it can be established that the evaluation system offers a fair picture of the quality of programmes. It functions as a control of programme quality.

Developmental aspects of the system are built into the evaluation model but can be seen most obviously in the follow-up phase. The effects are surprisingly large. It is likely that the results of the evaluations will be increasingly valuable for students choosing programmes but it is uncertain whether *Resultatsök* already fills this function at the moment.

Experiences unrelated to the system

The value of qualifications

Governmental control of higher education is exercised through the Qualifications Ordinance and qualitative targets. Higher education in Sweden is organised on the basis of *qualifications* that are subject to the Higher Education Act and the Higher Education Ordinance, above all to the Qualifications Ordinance. It is not, therefore, courses or even degree programmes that are the real object of evaluation (and even less the HEIs or their quality development procedures). The Higher Education Ordinance stipulates that “All first and second-cycle study programmes shall be offered in the form of courses. Courses may be combined to create study programmes.”(Higher Education Ordinance 6.13). This means therefore that courses are the basic units, whereas in international practice this is true of degree programmes. Moreover, in Sweden a larger proportion of higher education is offered in the form of free-standing courses than is the case internationally. A qualification can be attained either by completing a degree programme or by accumulating credits from free-standing courses. Both routes are subject to the qualification descriptors laid down in the Qualifications Ordinance (Higher Education Ordinance Annex 2).

The evaluation system is therefore based on *qualifications* as it is these that are evaluated. It has already been emphasised that one of the system’s strengths is that evaluation is based on targets whose validity cannot be questioned, while its weakness lies in the methodological difficulties of evaluating on the basis of targets. In its evaluations Sweden has a considerable advantage in the shape of the national targets laid down for higher education.⁵

⁵ In Australia TEQSA uses the same model for its evaluations. But as there are no national targets in Australia the entire evaluation process has to begin with discussions between the subject experts and the HEIs involved to establish which targets should apply for the evaluation of a specific subject area.

All over the world higher education is expanding and in Sweden it forms the largest state sector. The growth of the sector has led to the involvement of more stakeholders offering tertiary education. Today there is a large and ever growing market for education in which there are major variations in the form these offerings take as well as in their quality. Both large and small, serious and less serious education providers have been established. Some of the offerings are cheapjack while others provide highly advanced programmes.

MOOCs (Mass Open Online Courses) constitute a growing phenomenon in the world of education. One characteristic of these courses is that they can be scaled up almost without limit in terms of numbers of participants, that no prior qualifications are needed, that tuition is free of charge and that most of the “teaching” takes the form of web-based discussion between participants. The input from teachers is restricted and to some extent automated. MOOCs do not include examinations.

All of these new education providers compete *de facto* with higher education in Sweden today. But Swedish higher education has one advantage – entitlement to award qualifications. Academic qualifications are highly valued in the eyes of the public and the market in which graduates are employed. According to the annual surveys conducted by the SOM Institute, great confidence is placed in higher education by the general public in Sweden. Tertiary education can be offered by many agents today but only the HEIs can award qualifications when studies are completed. It is extremely important to safeguard the value of these qualifications. And this is exactly what the evaluation system does.

Self-contained courses

There is also a downside to focusing on qualifications and qualitative targets rather than courses. The Swedish system of higher education is based on the American “self-contained course” system, i.e. individual teachers are responsible for the teaching as well as the examinations in “their” courses. A qualification is awarded after completion of the courses required in a sequence that is sometimes laid down in a programme syllabus. If a course of study consists only of free-standing courses, it is ultimately the qualification requirements that determine whether a qualification can be awarded. Irrespective of how the courses have been completed, the qualification “comes into being” through the accumulation of the courses itself. On no occasion in the course of study is there any “final check” that the qualitative targets have been attained. This is presumed to be the consequence of adding the outcomes of one course to the next.

The evaluation system requires, however, some form of final product control as what it is seeking to determine is whether the qualitative targets for the course of study in its entirety have been attained. The evaluation system is not satisfied by the aggregation of successfully completed courses, nor, actually, is the Qualifications Ordinance either. Given the structure of higher education in Sweden, there is no system of checking the final product – the qualitative targets. The closest that can be found is programme syllabuses that lay down and guarantee that completion of a specified course of study will mean that all the qualitative targets will have been attained. There is no final test to demonstrate that this is in fact the case.

One risk with the evaluation system is that HEIs will tighten even further the selection and order of courses studied to provide guarantees that students attain the qualitative

targets. In practice this would lead to less freedom for the students and alternative courses of study would be restricted. This is not a desirable development.

The evaluation system is enforcing a system that demonstrates and documents the attainment of qualitative targets. But fundamentally the HEIs that award qualifications have the same need to document attainment of these targets. Developing a system of this kind is therefore a mutual concern for the HEIs and SHEA.

Final examinations and vivas used to exist in higher education in Sweden until the end of the last century, above all for official qualifications (roughly the same as today's professional qualifications). Final examinations could last for several days and include all the required courses in a degree programme. Final examinations were also common in many countries, above all in Central Europe, up until the Bologna reform. This reform required a restructuring of all programmes, which was undertaken in Sweden through the 1977 reform of higher education.

Qualitative targets

Focusing on qualifications and evaluating on the basis of qualitative targets provides greater understanding of these targets. The governance of higher education in Sweden had already been changed through the 1993 reform of higher education (the Autonomy reform) when a Qualifications Ordinance was introduced and management by objectives began to apply to higher education. The 2007 reform of higher education (the Bologna reform) brought with it a more explicit Qualifications Ordinance but management by objectives was already in place. Irrespective of the degree of detail in the Qualifications Ordinances, it was not until the 2007 reform that the sector had any clear understanding of the qualitative targets and their role in higher education. The evaluation system has acted as a catalyst for this process and enhanced the intentions of the Bologna reform.

The Bologna reform included the ideal of shifting focus from “teaching” to “learning”. The aim of education is not to teach – but for students to attain learning outcomes. In focusing so explicitly on the outcomes of programmes the evaluation system makes their targets clear and at the same time benefits student learning.

It is natural that the evaluation system aroused discussion of the qualitative targets. This has been noticeable in each panel of assessors and has spread to the entire field of study affected by each of the evaluations. Awareness of the targets means that an important discussion has been initiated in the sector about what the aims of higher education really are. In each programme corresponding discussion has started about which targets apply, if they are too vague, if they are appropriate, if they overlap and, above all, if they are pertinent. Discussion of this kind is salutary for all educational organisations.

When other stakeholders in higher education, such as employers', employees', professional and student organisations, have realised how the system of governance and evaluation work they have also shown interest in taking part in discussing the targets. This is, after all, how their own wishes and interests can be made known. As a result of this discussion the Government has declared its openness in principle to review the qualitative targets.

How many programmes are covered by the evaluation system?

When it was planned, the ambition was for the evaluation system to cover the majority of first and second-cycle programmes. The net has turned out to be coarser, however, than expected and a considerable number of these programmes will not be evaluated using this methodology.

Fewer than five independent projects

As has already been pointed out, evaluations of programmes with fewer than five independent projects have been excluded as this would provide a poor basis for qualitative assessment. A large number of programmes have turned out to be unable to fulfil this requirement, i.e. many higher education programmes are so small that they do not produce five independent projects in the course of three years. This occurs mainly in the second cycle. The list below presents the number and proportion of programmes excluded because of their size in rounds 1 and 2. The numbers excluded in later rounds were not as large.

Subject	Total no. of programmes	Excluded	Proportion(%)
Economic history	21	11	52
Work science	15	4	27
Informatics	38	8	21
Media & communication studies/journalism	46	9	20
Earth sciences	58	10	17
Sociology	49	7	14
Psychology	46	6	13
Economics	47	3	6
Political science	80	4	5
Business administration	75	1	1
Speech and language pathology	13	2	15
Gender studies	21	7	33
Occupational therapy	26	6	23
Pharmacy	26	5	19
Public health science	27	4	15
Law	27	4	15
Audiology	8	1	12
Physiotherapy	27	3	11

All of the programmes included in the evaluation project for health and medical care administration, biomedical laboratory science, orthopaedics, medical physics and optometry in round 2 had more than five independent projects.

The evaluation system reveals that higher education in Sweden is characterised by a large number of programmes that are small in the sense that few independent projects are produced. This probably means that few students take these programmes – and that few students are “affected” by the lack of evaluation of them. It should be possible to take

advantage of this awareness in planning future programmes and the number of places to be offered.

Each HEI makes its own decisions about how a main field of study is to be classified and therefore how large or small such a field is. In doing so the HEI also determines whether a programme is to be subject to evaluation or not. (See the section headed *Main field of study* on p. 33).

"Support subjects"

Some subjects/main fields of study can be characterised as support subjects in programmes in which they are not the main focus. Subjects of this kind are often studied by many students but few make them their main fields of study and therefore few write their independent projects in them. Subjects/main fields of study of this kind are Swedish, English, mathematics, statistics and some computer science. These can be very large subjects in the sense that they have large numbers of students but nevertheless they generate few independent projects. Nor do the independent projects produced necessarily reflect entire programmes. In the worst scenario a "large" programme (with many students) could be totally ignored in the evaluation system as too few independent projects are produced.

Higher education diplomas and short professional programmes

The course requirements for higher education diplomas and several short professional programmes include independent projects but there are no stipulations about how extensive they need be. As the evaluation system places great emphasis on assessing the independent projects and requires them to be comparable, it is not suitable for shorter programmes. Like its predecessor, the current evaluation system does not therefore cover two-year programmes leading to the award of either general higher education diplomas or those in the fine, applied and performing arts. SHEA will evaluate a selection of programmes of this kind using another system. This will be based on a model that focuses on outcomes in which the independent projects will be supplemented by other forms of input.

Newly established programmes

Newly established programmes that have not yet been able to produce any independent projects cannot be evaluated using a system that concentrates on these projects. These programmes include Master's programmes in economics and the extensive teacher education area, in which the first entitlements to award qualifications were granted in December 2010. Individual HEIs may also have established programmes that have not yet reached a level at which independent projects are produced. In these cases the HEIs are required to notify SHEA that the programmes should be excluded from the evaluation.

Programmes that are being phased out

Programmes for which the HEIs no longer intend to award qualifications are excluded from evaluation. In these cases a decision must have been made by the HEI on closure of the programmes concerned.

Third-cycle programmes

The instructions issued to SHEA also require it to evaluate programmes that lead to the award of third-cycle qualifications. The results of evaluation of third-cycle programmes will not, however, lead to any extra funding from the Government. The evaluation system that applies to first and second-cycle programmes is not suitable for use at third-cycle level as its focus on independent projects would overlap too much with the established form of final examination of doctoral theses by examination committees made up of external examiners. Development is still taking place of an evaluation system for third-cycle programmes. One possibility being investigated is potential coordination with the evaluation of research for which Vetenskapsrådet (the Research Council) is responsible.

Programmes that do not lead to the award of qualifications

The evaluation of programmes or courses that do not lead to the award of qualifications is arranged differently. In January 2010 the Government assigned the National Agency for Higher Education the task (U2009/427/UH) of determining which programmes did not lead to the award of a qualification and of continuously evaluating a selection of them on the basis of the requirements for higher education stipulated in Section 8–9 of Chapter 2 of the Higher Education Act. Three rounds of evaluations of randomly selected programmes have been carried out during 2011, 2012 and 2013. This type of evaluation will continue.

This means that the evaluation system has not covered as many *programmes* as anticipated but it is difficult to assess the total number that have not been included. At the same time it must be stressed that this is not in proportion to the number of *students*. In other words, the fact that a large number of programmes are not covered by the system only affects a small number of students. The system provides the important insight that there are probably too many small programmes and also that the next generation of the evaluation system must be adapted to the programmes actually offered in higher education in Sweden.

Main field of study

The Higher Education Ordinance lays down rules for *qualifications* and also states that there has to be a *main field of study* for general qualifications. There are no national stipulations about what constitutes a main field of study, however, and this is decided by each HEI. This means that there may be as many classifications of a specific main field of study as there are HEIs in Sweden. This also means that the name of a main field of study does not guarantee that programmes are the same or even comparable.

Lack of consistency about both what domain a main field of study should cover and what main fields of study with the same name do actually include gives scope for misunderstanding by both students and employers. It also creates a great deal of

additional work for SHEA in planning and carrying out evaluations of main fields of study and means that there is a risk that assessors with the wrong expertise will be recruited.

A large number of problems arise in connection with the different classifications of the main field of study in the evaluation system and this jeopardises the equivalence of the system. It has been pointed out above that differences in the range covered by a main field may lead to a “subject” not being evaluated at one HEI while at another it is – even though they are offered in the same way and in similar circumstances.

The predicament of the language subjects in higher education in Sweden has been recognised for many years. In recent years attempts to resolve the situation have consisted of creating larger families of language programmes and in this context the concept of a main field of study has been appropriate. The downside of this solution is that each language is still perceived as a unique “subject” even though it belongs to a wider main field of study. It has happened that in broadly defined main fields of study no independent project in an established “subject” (a language with few students) has been randomly selected even though it has been included in the overall evaluation. The language subjects also have to take the distinction between the linguistic and cultural science elements into account. This distinction is also dealt with differently in the classifications of main fields of study at the HEIs.

Practice when it comes to determining main fields of study varies – not only in the language subjects. Some HEIs prefer broad fields of study such as “medical science” and “chemistry” while others opt for narrower classifications and divide up these main fields into “anatomy”, “physiology”, “pharmacology”, “nutrition” in the first case or into “biochemistry”, “physical chemistry”, “analytical chemistry” in the second. There are even main fields of study that are defined differently in the first and second cycle at the same HEI.

In addition there are classifications of main fields of study that are doubtful from a disciplinary point of view, such as the inclusion of translation programmes in a language subject (the main field is extended) or a main field called medicine turns out to be a small section of an entire faculty domain (the main field is restricted).

This naturally brings us on to the familiar questions about how the concept of a *subject* evolves and how new subjects are established across previous subject boundaries. There is a major risk in attempting to set up central norms about what should constitute a subject or a main field of study, which is why the HEIs have been given the right to make their own classifications. But the weaknesses that ensue from local classifications of activities that are undertaken nationally, and internationally, become apparent in the evaluation system.

Double degrees

The Higher Education Ordinance list three types of qualifications – general qualifications, qualifications in the fine, applied and performing arts and professional qualifications. In distinguishing between these types of qualifications the Government has stated that they serve different purposes. In its bill on teacher education the

Government stated explicitly that a general qualification should not be awarded at the same time as a professional one (Bäst i klassen – en ny lärarutbildning [Top of the class – new teacher education]). Government Bill 2009/10:89, p. 13).

In higher education in Sweden students have to apply individually for the award of a qualification but their applications depend totally on how the HEI has structured their studies – often in the form of degree programmes – and how it informs students about which qualifications their studies will lead to.

It is up to each HEI, on the basis of the qualifications it is entitled to award, to issue qualifications to students on completion of their studies and fulfilment of the qualification requirements. SHEA's annual reports have shown that for some considerable time double degrees have been frequent. Of the 74,000 qualifications awarded in 2013 just over 10,000 were double degrees. This means that after completing their studies – usually a degree programme – students are more or less automatically awarded two qualifications. There are numerous examples such as the award of a professional qualification in biomedical laboratory science and a BSc in the same subject or both a professional qualification and a general degree in pharmacy. In the past this has also been true of qualifications in engineering or nursing. The Bologna reform in 2007 gave rise to a special debate about Master's degrees in engineering where some HEIs chose to base their programmes on general qualifications and others retained the focus on professional qualifications so that the possibility of acquiring double degrees became obvious, irrespective of the way courses were structured.

The evaluation system also reveals that double degree are frequent and give rise to special problems. One example of many can be found in the report of the panel of assessors on their evaluation of programmes in agriculture, forestry and horticulture, in which double degrees are mentioned specifically.

The qualifications system at the Swedish University of Agricultural Sciences (SUAS) is not easy to understand. Double degrees are listed for very similar or identical courses of study. Both the panel of assessors and the university's students find it difficult to grasp the system completely and it also varies from programme to programme. The panel finds it remarkable that the university can award a professional qualification in Forestry and a general MSc in the same subject for identical courses of study even though the qualitative targets for the two qualifications are patently different. The panel considers that this may mean that the programme is not optimised for either qualification and runs the risk of never gaining a higher evaluation than acceptable. This was confirmed by the outcome of the assessment of the degree projects. Some were well adapted to the targets for a general qualification, others for the professional qualification in forestry.

(Report from the panel of assessor 18 June 2013, p. 3)

There are even examples of HEIs that award triple degrees for one and the same programme. The panel of assessors in the evaluation of programmes in agriculture, forestry and horticulture identify the crucial problem with double degrees: the qualitative

targets for general and professional qualifications are different. In several evaluation projects a programme leading to the award of a professional qualification has been evaluated as having inadequate quality while the “same” programme was judged as having high quality if it led to the award of a general qualification. The reverse can also occur. The fact that this can happen illustrates that the real issue is that there are different sets of qualitative targets to be attained and that it is not certain that one and the same programme can fulfil both of them. On the contrary, in view of the stipulation in the Higher Education Ordinance that professional qualifications and general qualifications require different types of study programmes, it is unlikely that both sets of qualitative targets can be attained within the same programme. If it is possible and reasonable to do so, the need to differentiate between the two types of qualification should be questioned. The risks in trying to satisfy two sets of qualitative targets at the same time are, as the panel of assessors also observes, that neither set of targets is attained optimally.

Double degrees are also questionable on the grounds that they can mislead employers and the surrounding community.

The most obvious argument heard for double degrees is that international adaptation makes it necessary to describe Swedish professional qualifications in terms of traditional academic degrees as well. The background can be found, however, in the 1977 reform of higher education when a large number of professional programmes were incorporated into an integrated higher education system. Compared with other countries the resulting higher education system in Sweden was unusually comprehensive. The professional programmes that were incorporated into the university world naturally sought to adopt an academic structure. This found expression for instance, but far from solely, in organising programmes so that they would result in the award of a both a professional qualification and a traditional academic degree – often a Bachelor’s degree. This is the downside of the “academisation” of professional education. It has had a major impact on some programmes ever since 1977.

There are therefore powerful voices and some good arguments for double degrees. The evaluation system emphasises, however, that it is doubtful whether these arguments outweigh the shortcomings of such a system.

Calibration

As has been emphasised, this evaluation system is extensive and complex. Its ambition is to balance the quality of a programme in which business administration is the main field of study at the University College of Halmstad in 2011 against one in pharmacy at Uppsala University in 2012 and of a Master’s programme in engineering at Luleå University of Technology in 2013 against a degree in the fine, applied and performing arts specialising in graphic design at Konstfack (the University College of Art, Craft and Design) in 2014. For this reason one vital requirement is the calibration of all the evaluation projects with each other. The responsibility for this task must to all intents and purposes rest with SHEA.

In the first round there were considerable differences between the results of the projects. At one extreme 25 of the 67 programmes in business administration were evaluated as having inadequate quality (and only eight were judged to have very high quality), while a

mere 6 of the 47 programmes in earth sciences were given the same evaluation (and 16 judged to have very high quality). Sometimes the criticism has been levelled that the existence of such major differences between subjects is unlikely. It is claimed in other words that the stringency of the panels of assessors has varied in the different evaluations or target attainment measured using different parameters. Similar criticism has maintained that it would be unreasonable to believe that one-fifth of the higher education in Sweden is of inadequate quality – which is roughly the proportion evaluated in this way to date. This could then arouse suspicion that SHEA has not succeeded in providing all the panels of assessors with a shared and reasonable scale for their assessments.

The problem has declined, however. Gradually, as the evaluation system has included more and more programmes, an increasing number have been grouped in the middle – a shared norm has emerged.

When an evaluation project ends with a large proportion of the programmes being evaluated as having inadequate quality or attaining very high quality, any discussion has to start from the results themselves, i.e. the quality of business administration programmes is perhaps too low – when it is offered at 26 HEIs. And perhaps earth sciences does consistently attain high quality or very high quality in Sweden – which is what the Nordic assessors maintained.

It would, however, be presumptuous to believe that the evaluation system is totally fair in the sense that all Swedish programmes are measured against exactly the same scale. Using a peer-review system inevitably means that scope is provided for the different assessments of peers. The problems of calibrating expert opinions, assessment of applications for research funding, PhD examination committees, etc. are familiar. It is up to the authority responsible – in this case SHEA – to set out as clearly as possible the core levels, criteria and scale. Considerable, and relatively invisible, effort is devoted to this as well. It is more than likely that this calibration process by SHEA will be further augmented in the next generation of the evaluation system.

Transparent but mechanical system

Criticism was expressed of the previous evaluation system's lack of transparency in the decision-making process and that different scales of values were used for different evaluation projects. These defects are smaller in today's system, which is a transparent one. It is easy for external observers to trace how the final evaluation has emerged as all the material is readily available on the Internet. This material shows how each independent project has been assessed on a three-point scale in terms of each target, and sometimes sub-goal, and how the overall evaluation has been determined. It is also easy to compare all the programmes with each other – over time or by main field of study, professional qualification awarded or HEI.

The downside of the transparency of the evaluation system is its mechanical and accumulative features. This weakness derives from the way that the many qualitative targets, sub-goals and independent projects together with the three-point scale for each target or goal and the overall evaluation per target creates a multi-dimensional system with, at times, thousands of values. When all the values are weighed together there is a

risk that fixed formulas rather than academic judgement will produce the final evaluation. In that case no use is being made of the collective expertise of the subject specialists.

It goes without saying, however, that there is scope for qualitative assessments as well. This is found mainly in the individual judgements of each of the qualitative targets in each independent project. Only the experts, with the help of their academic experience and expertise, can assess whether a qualitative target in their own discipline or field of study has been attained irrespective of whether they do so on the basis of an independent project, a self-evaluation or some other input. The overall assessment of individual targets and entire programmes also allows scope for academic judgement.

Striking the required balance between the demands for transparency and comparability, on the one hand, and scope for academic and professional judgement on the other is not a unique problem for today's evaluation system. Evaluation always entails the risk that very explicit criteria can leave room for subtle and difficult accommodations that offer a false impression of objectivity. In weighing the mechanical aspects against the transparency they provide, a balance should be struck that allows greater scope for the overall assessments of the experts. A shift of this kind presupposes that it is acceptable to require less transparency and comparability.

ENQA

ENQA – the European Association for Quality Assurance in Higher Education – expressed criticism of the Swedish evaluation system in a report that was based on the application for membership submitted by the National Agency for Higher Education in 2011. One of the association's requirements is that each member has to submit to an appraisal process once every five years.

ENQA's evaluation is based mainly on what is known as the European Standard and Guidelines for Quality Assurance in the European Higher Education Area (ESG). These were adopted at a meeting about the Bologna process in Bergen in 2005 of ministers responsible for education and are currently being reviewed.

The criticism expressed by ENQA differs in character and in gravity. On some points it is unfounded – as when it claims that the evaluation criteria are not disclosed in advance. Other points are correct and will be remedied. This applies for instance to the ability of the system to generate useful information for the programmes and HEIs evaluated. As has been mentioned, all the stakeholders agree that reports should contain more detailed feedback even if it is difficult to satisfy ESG 2.1 fully when it comes to *recommendations* to the HEIs. But there is also criticism that is interesting in principle.

The central criticism from ENQA is that the Swedish evaluation system does not satisfy ESG 2.1. This requirement is worded as follows:

2.1 Use of internal quality assurance procedures:

External quality assurance procedures should take into account the effectiveness of the internal quality assurance processes described in Part 1 of the European Standards and Guidelines.

This requirement means that when assessing the operations of a HEI an external evaluation institution, such as SHEA, has to take the internal quality assurance system at the HEI into account.

1) The Swedish evaluation system assesses courses of study/programmes (or actually qualifications that the courses of study/programmes lead to) – not the HEIs. When Sweden introduced evaluation of higher education after the 1993 reform, it began by evaluating the HEIs' own quality development programmes. Gradually, with increasing "evaluation maturity" at Swedish HEIs, the focus shifted to evaluations of courses and programmes themselves. At the stage that applies today, the focus has shifted to educational outcomes. Since 2010 the HEIs have had full responsibility for their own quality assurance procedures and this is not evaluated by a central authority. This is exactly what ESG 2.1 demands.

2) Today's Swedish system places its focus on the outcomes of programmes and although the internal quality assurance procedures in the HEIs are a vital requirement if outcomes are to be of high quality, these procedures are not themselves assessed. In its decision on the National Agency's/Swedish Higher Education Authority's membership application ENQA writes:

The ENQA Board noticed the high relevance of the learning outcomes based approach for the recently introduced procedures of HSV, and believes in the importance to take into account the whole concept of learning outcomes in external quality assurance. However, the Board concluded that the ESG are sufficiently open to give space to various quality assurance approaches, and could easily accommodate the learning outcomes approach while at the same time respecting the standards and guidelines.

ENQA is saying, in other words, that the fact that the Swedish system focuses on results through its concern with learning outcomes should not prevent compliance with the demands in the ESG – including ESG 2.1. According to ENQA the demands in the ESG are open enough to allow a combination of ESG 2.1 and a focus on outcomes. It is difficult, however, to see how this would be possible. ESG 2.1 is a kind of preamble that lays down that a complete evaluation should be based to start with on a HEI's own quality assurance procedures. How this could be combined with the idea behind the Swedish system of disregarding the choice made by the HEI itself of the way in which it can attain the externally imposed targets in the Qualifications Ordinance is difficult to perceive. Here there is a difference in ideology. In the Swedish system the idea is to start from external criteria – the qualitative targets laid down in the Higher Education Ordinance – while ENQA's ideology is to start from the HEI's own internal quality procedures.

3) It can be seen from the description given in the section headed *A five-stage model* (p. 7) that the entire evaluation model starts with the quality development undertaken at the HEI itself and an application for entitlement to award a qualification. It is also made clear that each application of this kind means that SHEA carries out an evaluation of the HEI's own quality assurance system. This is a vital requirement if SHEA is to conclude that a high quality programme can be offered and entitlement to award a qualification granted. But it is not only in the framework of appraisal of entitlement to award qualifications that a HEI's internal quality assurance procedures are assessed in accordance with ESG 2.1.

It is possible for SHEA to include an assessment of the internal quality assurance procedures of a HEI in its current evaluation system but it remains unclear what role this would play in a system that fundamentally focuses on output – not input.

Resources

What does the evaluation system cost? In the national debate on evaluation of public services evaluation itself is often accepted but it is maintained that the balance between the resources devoted to it and the positive effects it can provide is uneven, i.e. that the costs of evaluation outweigh the benefits. Similar views are expressed in higher education. For this reason some factual calculations are valuable.

The total expenditure on first and second-cycle education amounted in 2012 to SEK 26 billion (State funding was capped at SEK 21 billion and in addition there were certain revenues. Årsrapport för universitet och högskolor [Annual Report on Higher Education Institutions]. SHEA Report 2013:2.)

SHEA allocates SEK 60 million each year to its quality assurance department. This is to fund both the appraisals of applications for entitlement to award qualifications and the evaluation system. In addition a small proportion is devoted to the evaluation of miscellaneous courses, for example, and specific assignments. About half of the resources allocated are used to remunerate the external assessors.

The proportion of the total expenditure on higher education in Sweden devoted to its evaluations by SHEA is therefore 2.3 per mille.

As has been pointed out the bulk of the work on quality assurance in the higher education sector is carried out by the HEIs. It is very difficult to calculate what resources are devoted to this work. Chalmers University of Technology has however published a report that reveals its expenditure for activities linked to the national evaluation system. (Utvärdering av Chalmers process för leverans av underlag till nationell utvärdering av examina 2012 (HSV-utvärderingen) – Slutrapport [Evaluation of procedures at Chalmers to provide material for the national evaluation of qualifications in 2012 (HSV evaluation) – Final report]. Reg. no. C2011/1196. Published 29 April 2013.)

Chalmers estimates that 26,000 working hours were devoted to dealing with the evaluation system for programmes in engineering in 2012–2013. Given an average annual workload of 1,730 hours this amounts to 15 man years. In 2012 Chalmers had a staff of 1,951 full-time equivalents. This means that in one year Chalmers devoted 7.7 per

mille of its man years to the national evaluation system. As evaluations take place at four-year intervals the resources allocated per year is about 1.9 per mille man years. Chalmers has a high proportion of researchers on its staff. If the calculation is made for teaching and administrative staff alone, the proportion of the total number of man years allocated is somewhat higher.

There may be other expenditure than that reported here by the HEIs. The cost of evaluation should be balanced against its positive impact on the quality of programmes. This is, however, virtually impossible to calculate. Nevertheless, on the whole it can be seen that the resources required by the evaluation system are small.

Looking ahead

Work has begun at SHEA on the development of the next generation of the evaluation system that is to apply from 2015.⁶ Future systems will be based on two premises that already apply in the current system.

- Focus on outcomes
- Evaluation on at least a three-level scale

The concentration on qualifications will be retained. As the future evaluation system will also be based on peer review, close cooperation with representatives of the sector is taken for granted. The development work will be undertaken in collaboration with stakeholders such as SUHF (The Association of Swedish Higher Education), the Swedish National Union of Students and representatives from the labour market.

A large number of the parameters in the evaluation system have been analysed above. In some cases review of a problem has concluded with suggestions for development. These will be summarised in this section and some new ideas introduced. These potential developments differ in status – some have almost been decided on while others are drafts for proposals – which means that these are not the only ideas about what may apply in the future.

Development idea 1: Broader results

In the section headed *Basing evaluations on the independent projects* (p. 19) the question was raised of what “outcomes” should be taken to mean. There has been the misunderstanding that these only comprise the independent projects. This has never been the case but the evaluation projects in the first round were deliberately selected to include programmes in which there is a strong tradition and acceptance of independent projects. Subsequently, as more professional qualifications have been evaluated, the self-evaluations and to some extent other forms of input as well have been given greater importance.

⁶ In April 2014 the Government appointed a special commissioner whose task was to propose an evaluation system from 2015.

The self-evaluations also focus on accounting for outcomes or explaining them. As has been pointed out, the likelihood of certain outcomes has to be confirmed by describing various types of processes or policy documents. It is very difficult to present “outcomes”, in the strict meaning of the term, that show that all qualitative targets have been attained.

In developing the next generation of the evaluation system the work of determining what material can demonstrate the attainment of qualitative targets should be extended and systematised. A shared and uniform conceptual system should be developed. It is particularly important to improve the selection of material that can justify a very high quality evaluation, as the value of the sources provided by the current input is limited. Greater acceptance of the system can be achieved by basing assessment of an entire programme on more extensive material.

Development idea 2: A greater wealth of information

The section headed “*Surplus information*” (p. 24) contains a discussion of how much information about individual programmes and the field of study evaluated is included in the system and the desirability of communicating it to the HEIs. It is a waste of resources not to make use of information that is available. The international practice in evaluating higher education is to feed back as much information as possible to the HEIs and this usually takes the form of recommendations about how programmes should be improved. Lack of feedback is one of the points raised by ENQA in its assessment of the Swedish evaluation system. The HEIs ask for more detailed information about shortcomings so that they can make appropriate improvements.

At the same time the fundamental ideology of the system must not be questioned: the evaluation by the central authority must not have the effect of setting norms – the HEIs must analyse their shortcomings and merits autonomously and take measures based on this kind of analysis.

In the next generation of the evaluation system a new balance should be struck between these two aspirations. It should be possible for a final report from a panel of assessors to include a more detailed description of what is considered problematic in a specific programme. It should also be possible to describe possible solutions without recommending which of them should be used. And also with regard to the “state of the art” in a subject or programme it should be possible to draw general conclusions about their national standing and propose improvements to a much greater extent than today. This would also alter the balance between control and development – control would still be explicit but the developmental features would be much more obvious.

Developmental idea 3: More scope for the experts

In the section headed *Transparent but mechanical system* (p. 37) the balance between the transparency of the system and the scope for peer judgements was discussed. The world around us is demanding more and more from all evaluation systems and the more stringent the sanctions linked to their results become, the greater the demand for transparency and impartiality. Today’s system has a high level of transparency but also includes a mechanical element that can give an inviting impression of objectivity. The

system is also large and complex and this can increase the risk of resorting to simplistic arithmetical calculations.

Ultimately, however, the system stands and falls on the subject experts' qualified assessments of whether qualitative targets have been attained. This kind of judgement can only be made by a subject expert or an expert in a specific field of study and the system is therefore based fundamentally on peer review.

The element of peer review will obviously characterise future evaluation systems and consideration should be given to the possibility of enhancing it. This could take the form, for instance, of providing greater scope for a summary in the final evaluation in which the assessors could share the knowledge and the understanding they have gained of how the subject is taught with their peers in an overall assessment of all the material. Adjusting the balance between the requirement for transparency and scope for peer assessments is something that would have to be discussed with representatives of the HEIs.

Development idea 4: Evaluate the HEIs' quality assurance system

Most of the work on quality assurance in higher education is undertaken at the HEIs. This work is a prerequisite if a HEI is at all to be able to apply for and be granted entitlement to award a qualification and if it is going to be able to offer high quality programmes. In the previous system the actual quality assurance procedures at the HEIs have been the subject of the national evaluations. This system was criticised by the student organisations, among others, who claimed that it was what students learnt that determined whether the quality of a programme was high, not the quality of the quality assurance system. The subsequent national evaluation system therefore focused on the quality of programmes by evaluating their circumstances, processes and outcomes. The system used today focuses on outcomes.

ENQA maintains that ESG 2.1 requires the HEIs' own quality assurance systems to be the starting point for national evaluation. SHEA conducts this kind of evaluation in the framework of its appraisal of applications for entitlement to award qualifications.

Can evaluation of the HEIs' quality assurance procedures be integrated into the next generation of the evaluation system – which will also continue to focus on outcomes?

Development idea: 5 More explicit labour market preparation

All education is for use – in one way or another – and a great deal of education offers preparation for the labour market. The Higher Education Act states – somewhat cryptically –

First-cycle courses and study programmes shall develop ... the preparedness of students to deal with changes in working life.

(Higher Education Act Chapter 1 Section 8)

Evaluation is made on the basis of the targets laid down in the Qualifications Ordinance and these do not include a labour market perspective as explicitly. A reasonable

interpretation is that the accumulation of all the qualitative targets will result in the student having a high degree of preparedness to deal with changes in working life. In the guidelines on the selection of targets it is stipulated that the labour market perspective is to be taken into account.

Preparation for the labour market should be given a more explicit role in the next generation of the evaluation system. What is expected of the labour market representatives in the panels of assessors should be expressed more clearly.

Development idea 6: Minimise homogenisation effects

All evaluation has levelling effects. In the worst cases an evaluation system can pervert the subjects of its evaluations so that they adapt what they do solely on the basis of the indicators used. The current system has been designed within the framework of a broader ideological assumption that higher education and research have greater potential for development the more autonomous the HEIs are. The evaluation system has been shaped to provide a counterpart to the increasing autonomy of the HEIs. The focus on outcomes means that only the national targets are to govern higher education while how these targets are attained – teaching staff, study environment, course syllabuses, laboratory equipment, quality assurance systems, etc. – are the responsibility of the HEIs. In this way diversity can be encouraged in higher education.

There are preliminary indications that the system is not capable of living up to its ambition to encourage diversity. There is a risk that the variety of study approaches is being limited and that the structure of the independent projects standardised. In all educational systems interdisciplinary courses of study pose problems, such as entrepreneurship programmes, and innovative ones that are not assessed correctly.

Analysis should be made before the next generation of the evaluation system on what homogenisation effects the current system has and how they can be minimised.

Development idea 7: Augment calibration

The section headed *Calibration* (p. 36) discussed the problem of calibrating a very large number of evaluations over several years that cover all subjects and programmes. The ambition to place all higher education in Sweden in a joint evaluation framework is both very high and correspondingly difficult. The responsibility for calibration must rest with SHEA.

Before the next generation of the evaluation system calibration between different evaluation projects must be enhanced, above all through providing the panels of assessors with more explicit guidelines.

Development idea 8: Enhance the student perspective

Student perspectives are included in the evaluation process both by engaging student as members of the panels of assessors and also through the input provided by student interviews.

The shift to a focus on outcomes in the system has presented a challenge when it comes to including the students and their perspective in the evaluation process. Fundamentally this is because of the focus on the independent projects but the challenge is also a result of the difficulties that students have in accounting for – and assessing – the outcomes of a programme they have not yet completed. Inclusion of the students' perspective in the evaluation process is not, however, controversial so development of the student perspective will be intensified.

What is most important is to make the role of the student assessors explicit. How students are expected to contribute to the assessment of the quality of a programme must be clarified. This can be done by giving students concrete tasks in the evaluation process. Further development of the different forms of input and how they are dealt with will take student perspectives and participation into account. In addition students should be given more support in their role as assessors. A clear training programme should be developed. Finally the administration of the student interviews should be improved.

Concluding reflections

SHEA's evaluations form part of the way in which the Government manages higher education. Here higher education is not a special case – quite the opposite, as management by goals and outcomes is today the predominant form of governance in the public sector. This management is intended to ensure that the objectives of public service are attained and to improve its outcomes. Stipulating targets and evaluating results is beguilingly simple and it is difficult to maintain that it is wrong. Extensive discussion is taking place about New Public Management – the context in which management by goals and outcomes prevails – and criticism is expressed. The crux of this criticism is that evaluation systems tend to measure what can be measured and the subject of an evaluation must make itself "evaluable". The result is that the evaluation system subverts organisations instead of contributing to their development.

In this context, the current Swedish system for evaluating higher education possesses considerable advantages. Because the Qualifications Ordinance contains explicit national targets for higher education there is no need to make evaluation possible by constructing retrospective objectives on which it can be based. Nor does the evaluation system confine itself to measuring what can be measured. On the contrary – the qualitative targets are difficult to evaluate and therefore peer review is used throughout to assess whether they have been attained. It would, of course, have been much simpler to measure student completion rates, the number of teacher-led contact hours, number of qualifications awarded, per capita costs, etc. In other words – the evaluation of higher education has a major advantage in being based on national targets and in applying peer review.

Another criticism is that evaluation that lays stress on control gives rise to distrust and undermines the roles of professionals. Academic teachers are highly professional but are nevertheless not immune to this risk. It is vital for the Swedish evaluation system to develop in ways that do not inhibit the natural desire of academic teachers to develop.

On the other hand evaluation of higher education cannot avoid the general problems linked to the process involving its retroactivity, costs, methods of selection and measurement problems.

Education can be evaluated for different purposes, for different stakeholders and in different ways. Basing evaluation on the current qualitative targets is a reasonable purpose that above all satisfies the interests of the Government. There are alternative legitimate reasons and interests. The methodology applied in today's system is adapted to its purpose and the public interest. Educational quality has been defined operationally – degree of attainment of the educational targets – a reasonable definition given the guarantee needed by the Government that stipulated targets are attained.

The methodology is based on gathering information that is likely to demonstrate target attainment. Here there is undoubtedly room for improvement. More material needs to be collected and analysed systematically, more information can be extracted from the system, the peer-review element should be enhanced, calibration improved, etc. There are also details that can be adjusted. Consideration should be given to whether zero tolerance should continue to apply to attainment of the qualitative targets, whether the relative weighting given to the different forms of input should be reconsidered, whether more assessors should read the independent projects – at least in doubtful cases, whether the threshold for the number of targets that have to be attained for an evaluation of high quality be altered, whether another terminology should be used to express evaluations, etc.

These methodological problems are real ones and will be dealt with, but their existence does not give any grounds for questioning the Swedish evaluation model. Quite the contrary: it is a model that functions well, is based on the intentions of the Bologna reform and sets a good example for Europe.

The Swedish Higher Education Authority is a government agency in the higher education sector. The Authority evaluates the quality of higher education, exercises legal supervision, reviews efficiency and is responsible for statistics and monitoring the sector.

www.uka.se

