# Conformity or confusion? Changing higher education grading scales as part of the Bologna Process

*The cases of Denmark, Norway, and Sweden*

Dahl, Bettina; Lien, Eirik; Lindberg-Sand, Åsa

*Published in:*
Learning and Teaching: The International Journal of Higher Education in the Social Sciences

*DOI (link to publication from Publisher):*
10.3167/latiss.2009.020103

*Publication date:*
2009

*Document Version*
Publisher's PDF, also known as Version of record

in Higher Education, < http://depts.washington.edu/cirgeweb/c/research/social-science-phds%E2%80%94five-years-out/ > (accessed 15 December 2008).

No Child Left Behind Act (2001) 'Public Law 07-110, the No Child Left Behind Act of 2001', < http://www.ed.gov/policy/elsec/leg/esea02/index.html > (accessed 15 December 2008).

Newfield, C. (2004) *Ivy and Industry: Business and the Making of the American University, 1880–1980*, Durham: Duke University Press.

Polanyi, K. (1944) *The Great Transformation: The Political and Economic Origins of Our Time*, New York: Rinehart & Co.

Report of the Spellings Commission (2006) *A Test of Leadership: Charting the Future of U.S. Higher Education*, Washington, D.C.: U.S. Department of Education, < http://www.ed.gov/about/bdscomm/list/hiedfuture/reports/final-report.pdf > (accessed 15 December 2008).

Ruben, B., Lewis, L., Sandmyer, L., Russ, T., Simulowitz, S. and Immordino, K. (2008) *Assessing the Impact of the Spellings Commission: The Message, the Messenger, and the Dynamics of Change in Higher Education*, < http://www.nacubo.org/documents/business_topics/AssessingImpact.pdf > (accessed 22 December 2008).

Spellings, M. (2006) 'Charter for a national dialogue: The Secretary of Education's Commission on the Future of Higher Education', in *Report of the Spellings Commission, A Test of Leadership: Charting the Future of U.S. Higher Education* Washington, D.C.: U.S. Department of Education, < http://www.ed.gov/about/bdscomm/list/hiedfuture/reports/final-report.pdf > (accessed 15 December 2008).

Weick, K. (1976) 'Educational organizations as loosely coupled systems', *Administrative Science Quarterly* 21: 1–19.



# Conformity or confusion? Changing higher education grading scales as a part of the Bologna Process: the cases of Denmark, Norway and Sweden

*BETTINA DAHL, EIRIK LIEN AND ÅSA LINDBERG-SAND*

**ABSTRACT**

The aim of the Bologna Process is to make higher education systems across Europe more transparent. It is crucial for this purpose that confusion concerning the characteristics of the systems should be replaced by conformity. But, as we will show, conformity brought about at one level may create confusion at another. The curricular aspect of the Bologna Process focuses on a shift to outcome-based and student-centred programmes. Syllabi should now be based on intended learning outcomes (ILOs) and should be adjusted to general level descriptors for qualifications. However, the Bologna documents give no explicit recommendations about the use of grading scales. In Denmark, Norway and Sweden, the reforms of higher education induced by the Bologna process included a change of grading scales and referred to the European Credit Transfer and Accumulation System (ECTS). Through these three case studies, we describe and analyse the political process and argumentation underpinning the decisions to change the grading scales in each country. This includes the problems, both experienced and perceived, with the old grading scales, the various national assessment traditions and the new grading scales. The purpose of the change was not the same in each country, but the ongoing adaptation to a seven-step grading scale was thought to ease the international recognition of the national grades, making mobility easier. Though a seven-step grading scale was implemented in both Danish and Norwegian higher education and also by an increasing number of Swedish higher education institutions, the translation of grades only works on a superficial level. The grading scales designed are fundamentally different as classification systems; they attach different numerical values to grades with identical labels and they relate differently to norm- and standards-referenced judgements of learning outcomes. The information condensed in similar grades from the three countries cannot be

*Learning and Teaching*
doi:10.3167/latiss.2009.020103

Volume 2, Issue 1, Spring 2009: 39–79 © Berghahn Journals
ISSN 1755-2273 (Print), ISSN 1755-2281 (Online)

equated. The vision of simple transparency turns out to be an illusion.

KEYWORDS
Bologna Process, European Credit Transfer and Accumulation System, grading systems, higher education, intended learning outcomes, Scandinavia

## Introduction

The aim of the Bologna process is to bring higher education systems across Europe into alignment. One of the tools for achieving this is making grading scales comparable. Some countries have reformed their higher education grading scales with the ostensible purpose of having systems that students, universities, parents and employers will understand anywhere in Europe. Our[1] study explores the reforms of the grading scales in three Scandinavian countries, Denmark, Norway and Sweden. These three countries are chosen because, of any European country, they are the most alike, with shared histories, similar cultures and higher education systems that are very close to each other. By studying the attempt of three countries, whose systems were already closely aligned, to use the Bologna process to try and make their grading systems comparable, we show that the result is an appearance of similarity that cloaks substantial and important differences. Now, however, it is more difficult to fathom and interpret the significance of those differences. This detailed study of attempts to create comparability in grading scales raises questions about whether other aspects of the Bologna process lend themselves to creating a fiction of a 'higher education area' in Europe.

The Social Democracy Party in each country has had a dominant influence on the development of the countries' quite similar welfare systems, so that they are known collectively as the *Nordic (Social Democratic* or *Institutional)* welfare states (Bryson 1992: 110–119). These Scandinavian countries also have a common past, an almost identical language and quite similar cultures. During the last decades, the Nordic model has been challenged by several factors. Denmark and





Sweden are members of the European Union (EU), while Norway is not. Politically, all the countries have changed regimes more often than before, swinging between social democratic and conservative/liberal[2] governments. In terms of the education sector, the Scandinavian countries also have rather similar education policy objectives, which are to secure equal access to learning, teach democracy, foster independence and critical awareness in pupils and increase equality. The focus is broad, comprehensive and non-elitist. However, there are also differences in the organisation and level of centralisation of the education systems; differences that also characterise the organisation of higher education, including the assessment systems and grading traditions (Dahl 2003; Fägerlind & Strömquist 2004). This will be elaborated and discussed below.

The three countries have all referred to the Bologna Process as a reason to change the grading scales used in higher education to (mostly) a seven-step model with two failing grades. In doing so, they all refer to the scale used for translation of grades included in the European Credit Transfer System (so called ECTS grades). If transparency or even a harmonisation concerning grading in European higher education should be regarded as a possibility, as discussed by Karran (2005), these three countries would be a good example to study. The countries are culturally quite similar, when compared to the remaining 43 nations also included in the Bologna Process, and they have all chosen to base their new grading scales on the same model. But our comparison will reveal that important differences are nonetheless hidden beneath the surface of coordination, calling into question the promise of a harmonised system.

We have studied both the process through which the grading systems changed as well as the product, i.e., the new scales. This includes studying differences in national and institutional controls of higher education assessment systems. We assumed that these differences emerge and become visible particularly when major structural changes such as a new grading scale is developed and implemented. Our analysis is what is often called a 'study of policy content'. Hence, we describe the genesis and development of the new grading scales,

how the policy emerged and give an overview of how it was implemented. This approach is combined with a view on grading scales as a special kind of classification system (Bowker & Star 1999). A classification system is a human artefact, both material and symbolic in nature. Bowker and Star (1999: 10) define a classification system as a set of boxes into which things can be put, to then do some kind of work. The classification itself is a spatial, temporal or spatial-temporal segmentation of a part of the world and/or of human knowledge. For instance, the division of a day and a night into 24 hours is a simple and global classification of time. We are deeply dependent on the way classification systems work, as they shape a great part of the taken-for-granted infrastructure of our civilised lives. This is especially so if the systems, through political or technical decisions, are made standard for social practice. Widespread standards soon naturalise and become invisible preconditions for ordinary situations. When classification systems are changed or replaced, they reappear on the agenda. Classifications work as objects for co-ordination and co-operation across social worlds. A classification may acquire the property of a boundary object. Such an object has a certain design that is meant to be identical when it is reproduced in different communities of practice. But when it is brought to the fore or utilised in different communities, the meaning and function of the object may be changed, though its design is still the same. If so, a classification system may only *appear* to create a common standard. And though the classification system may still work as a link between the social practices in question, its effects, compared to the intentions with the system, will be partly illusory and partly hidden. The Bologna Process builds on the development and acceptance of several new classification systems covering higher education cycles, qualifications, credits and grades.

Thus, our focus is not the policy process in itself, i.e., all the stages it went through, or which actors influenced exactly what. Instead, we focus on the design and implementation of the grading scale as a crossroads between (a) policy content, (b) national regulations of assessment and (c) the specific grading scale as a quite rigid artefact, a classification system used as a tool for policy expectations and



implementation at several levels (Bowker & Star 1999; Ham & Hill 1993: 9–10).

## The Bologna Process and the ECTS

### *The Bologna Process*

In 1999, the Bologna Declaration was signed by the education ministers from 29 European countries and now 46 countries altogether have signed the agreement. It is a commitment voluntarily accepted by each country to reform its higher education system to create overall convergence at European level in order, for instance, to increase the mobility of students, teachers and researchers (Dahl 2003). To create a mobile European Higher Education Area (EHEA) by 2010, the Bologna Process puts priority on the following three instruments (EU 2008a):

1. Introduction of the three-cycle system (bachelor/master/doctorate).
2. Quality assurance.
3. Recognition of qualifications and periods of study.

From 2003, and later confirmed by the Ministerial meeting in Bergen 2005, the Bologna Process has increasingly developed a focus on curricular reform. The Bologna Stocktaking Process Report 2007 emphasised that all the aspects of the Bologna Process are interdependent and linked to each other by the shift to a curriculum based on learning outcomes:

> There are two themes that link all the action lines: a focus on *learners*, and a focus on *learning outcomes*. If the Bologna Process is to be successful in meeting the needs and expectations of learners, all countries need to use learning outcomes as a basis for their national qualifications frameworks, systems for credit transfer and accumulation, the diploma supplement, recognition of prior learning and quality assurance (Bologna Stocktaking Process Report 2007: 3).

We would argue that this might *appear* to be a simple matter of a common 'currency'; however, it is a much deeper reform of the educational strategy and involves the acceptance of an increasingly uniform curricular system.

The Bologna Process is not an EU initiative, but the commitments, goals and instruments fit a lot of the EU initiatives, such as the Copenhagen Declaration of 2002 that had the purpose of enhancing co-operation in European vocational education and training (VET). In 2006, the EU Commission proposed the European Qualifications Framework for lifelong learning (EQF), which is supported by the EUROPASS (documents to make a person's skills and qualifications clearly and easily understood across Europe), credit transfer (ECTS-ECVET) and quality assurance (ENQA-ENQAVET) initiatives. The European Qualifications Framework for lifelong learning was approved by both the European Parliament and by the Council, and it was published in April 2008. The framework should be used as a translation grid for qualifications all over Europe. It is firmly based on the view that the results of human learning can be described as *learning outcomes*, qualitatively described in three different forms: knowledge, skills and competence. The framework includes eight levels of qualifications (expressed as knowledge, skills and competence) covering lifelong learning and corresponding to the whole educational system, from compulsory schooling to a research degree. Levels 6, 7 and 8 in the European Qualifications Framework correspond to the three cycles (bachelor, master, doctorate) in the Framework for Qualifications in the European Higher Education Area (EQF-HE), which is part of the Bologna Process. The latter framework was approved by the ministerial summit in Bergen 2005 (BWG-QF 2005) and is now a central tool for co-ordination in the Bologna Process. The framework has the character of an overriding meta-curricular framework, including level descriptors in the form of general *intended learning outcomes* (ILOs) for each of the three cycles. We would argue that if the agreements concerning the EQF and EQF-HE are not incorporated into local syllabi and realised in the teaching and learning processes, they will not affect the actual educational quality. According to the Standards and Guidelines for Quality Assurance in the European Higher Education Area (ENQA 2005), student



assessment should examine the achievement of the ILOs, which would imply a grading system based on standards – or criterion-referenced assessment, also called an absolute grading. And this is the point at which the new frameworks and older grading systems are not fully compatible, since, for instance, the Scandinavian grading scales were not constructed to assess intended learning outcomes even though the scales were used for criterion-referenced assessment. But how could and should grades be understood and compared in the European Higher Education Area? When learning outcomes are used as the foundation for describing both the goals for and the qualifications acquired in higher education, as stated by ENQA above, the question of how student assessment and grading are carried out is brought to the fore. This is also the point at which we see a similarity between the development of a European Higher Education Area initiated by the Bologna Process and the EU creation of a common market. Are grades suited to become a common currency in the European Higher Education Area?

### *ECTS and different types of assessment*

The reason behind the development of the ECTS scale was that the great variations in European grading scales made comparisons between grades from different countries more or less impossible, since the scales were constructed and used in different ways (Karran 2005). One of the problems was that in some countries the grading scale had an impressive number of levels, but in practice only the top grades were used and assigned to the majority of students. The ECTS scale was designed to solve this problem by applying Norm-Referenced Assessment (NRA) by stipulating that the ECTS grade should show if the original grade was issued for a minor proportion of the students with top performances or for a greater group of students with average performances. This is also called a relative scale. Norm-Referenced Assessment aims at telling which students performed better than others. It is based on the assumption that human traits and abilities are regularly distributed over a population as seen, for example, when individual differences concerning stable characteristics in psychology

are tested. This is a contested assumption in relation to study efforts and learning, where several other aspects are more relevant to consider when grading students. However, Norm-Referenced Assessment often stipulates that there should be a constant proportion of students in each grade awarded. To work as stipulated, the grading has to be linked to an assessment system where big groups of students are given the same assessments. Norm-Referenced Assessment is different from Criterion-Referenced Assessment (CRA) where the grading of students tells what the student has learned and how well. The grade awarded is thus independent of other students' grades. It is based on the standards (criteria) for student learning as course or unit objectives or Intended Learning Outcomes (ILOs). If the Intended Learning Outcomes are arranged in a hierarchy with different levels, the student's performance can be graded according to the level he/she is performing at after a qualitative assessment. In a Criterion-Referenced Assessment, distributions of grades cannot be stipulated beforehand. The relation between Norm-Referenced Assessment, Criterion-Referenced Assessment and the different grading scales that are used is not at all clear-cut. A grading scale may be more or less related to the assessment system, some grading scales might even be used for both types of assessment systems. This may be used as an example of the fact that the symbolic value attached to a classification (in this case a grade) or the actual usefulness of the classification in society may be regarded as more important than the coherence of or validity in the content that is classified by the system. Bowker and Star (1999) give several examples of classification systems of this kind; for instance, systems for medical diagnoses.

The European Credit Transfer and Accumulation System (ECTS) was developed at the beginning of the 1990s as a tool for increasing institutional transparency in European programmes for student exchange. Aspects of the ECTS are now adopted and made part of the curricular reform in the Bologna Process. The most important of those aspects is the quantitative measure for student workload as the foundations for assigning a certain amount of credits to a course or a programme, where 60 ECTS credits is equivalent to one-year full-time



study. In this paper, though, we focus on the part of the ECTS that is described as a kind of grading system. Here, after the student receives the national grades, ECTS grades are assigned among the passing students on a statistical basis as follows (EU 2008b):

A: best 10 %
B: next 25%
C: next 30%
D: next 25%
E: next 10%

There are two failing grades: 'Fx' means 'some more work required to pass' and 'F' means 'considerable further work required to pass'. The ECTS grade is not meant to replace the original national grade. It is produced as an administrative measure only to explain it, hence it has no legal value of its own. The ECTS scale is supposed to be only a translation tool, not a 'real assessment'. But the intriguing question is: what are the consequences for student assessment when one classification system based on an NRA perspective (the ECTS translation tool) is put on top of another one (the original grading scale, based either in CRA or NRA)? When the old classification is re-classified – who is ruling? In the London Communiqué of 2007, the Ministers stated:

> We underline the importance of curricula reform leading to qualifications better suited both to the needs of the labour market and to further study. Efforts should concentrate in future on removing barriers to access and progression between cycles and on proper implementation of ECTS based on learning outcomes and student workload (London Communiqué 2007: 2).

The last sentence above implies the focus of the curricular reform; all European higher education syllabi in the future should be based on learning outcomes. If implemented, it will lead to a shift with huge educational consequences, including at the classroom level. That a 'proper implementation of ECTS' should include a reform of national or institutional grading scales is not mentioned anywhere in the Bologna documents. When problems with the implementation are

discussed, they are mostly concerned with the introduction of national credit systems or the shift to describe the scope of a module or a programme in terms of calculated student workload instead of teaching hours. The agreement on the Framework for Qualifications in the European Higher Education Area (EQF-HE) and the European Qualifications Framework (EQF), both presuppose that qualifications and degrees are awarded to students who fulfil the Intended Learning Outcomes, and the Standards and Guidelines for Quality Assurance in the European Higher Education Area (ENQA 2005) states that student assessment should examine the achievement of the Intended Learning Outcomes. This could imply a recommendation to use Criterion-Referenced Assessment as a base for grading, but this is not explicitly stated. The ECTS grades are, however, firmly based in Norm-Referenced Assessment. Thus, there may be a tension growing between the educational perspective on student assessment expressed on the one hand in the new frameworks for qualifications and on the other hand in the translation of grades in the European Credit Transfer and Accumulation System. Still, Denmark, Norway and Sweden referred to the Bologna Process when starting to implement ECTS-like, absolute grading scales, which we will discuss below.

## Denmark

### *Higher Education in Denmark*

The Ministry of Education covers primary, secondary and tertiary education (e.g., teacher training colleges), while universities and research institutions are part of the Ministry of Science, Technology and Innovation. It is a binary system where vocational programmes are not part of the university system. The Danish university sector has experienced a number of major changes in the 2000s. The new Danish liberal/conservative government from 2001 had a University Law approved with support from the Social Democratic party and the Christian People's Party in 2003, which made the Danish universities self-governing institutions with boards as the top authority. Board members from outside the university form the majority, and the chairman of the board must also come from outside the university. The rationale for the new law was the rapid change in the demands that society made of the universities. Not only had the organisation of the universities changed, but also their numbers. In 2006, Denmark had 12 universities and 15 research institutions. All universities are mainly publicly funded whereas the research institutions have a mix of public and private funding. The Minister of Science, Technology and Innovation wanted to reduce this number to strengthen Danish research and teaching internationally, enhance co-operation with business life, make it easier to attract EU funding and make it more possible to do commissioned work. After a process of mainly voluntary fusions, the result in 2007 was eight universities and three research institutions (Danish Ministry of Science, Technology and Innovation 2005, 2008).

### *The old and the new Danish grading scales*

The former Danish grading scale was called the '13-scale' (see Table 1). It was constructed as a Norm-Referenced Assessment system – relative scale – and had been the national grading scale used in com-

**Table 1 The former Danish grading scale, the 13-scale (Official translation from Danish)**

| Old Grade | Description |
|---|---|
| 13 | Is given for the exceptionally independent and excellent performance. |
| 11 | Is given for the independent and excellent performance. |
| 10 | Is given for the excellent but not particularly independent performance. |
| 9 | Is given for the good performance, a little above average. |
| 8 | Is given for the average performance. |
| 7 | Is given for the mediocre performance, slightly below average. |
| 6 | Is given for the just acceptable performance. |
| 5 | Is given for the hesitant and not satisfactory performance. |
| 03 | Is given for the very hesitant, very insufficient and unsatisfactory performance. |
| 00 | Is given for the completely unacceptable performance. |

pulsory education (grades 1–9) and high schools/*gymnasium* (grades 10–12) since 1963. In 1971, higher education, including universities, was required to also use the 13-scale, but now as a Criterion-Referenced Assessment system – absolute scale. In 1992, grades in high school were also changed to CRA, and finally, in 2000, compulsory education, where grading begins in grade 8, was also required to use the 13-scale as an absolute scale.

Parallel to this scale, a grading scale of pass–fail was also used. At the universities a maximum of one-third of the grades could be pass/fail, whereas these grades were used much more in, for instance, nursing and midwife education to avoid competition. When grading, examiners first decided on the overall category: fail (0–3–5), pass (6), middle (7–8–9) and high (10–11–13); then the exact grade was decided. The 13-scale is much in line with Bloom's Taxonomy (Leth Andersen 2005) and most often, to get more than 9, one needed much more than 'just' to know the material. Furthermore, when the scale was used as a Norm-Referenced Assessment system, 8 had to be the national average for each subject at a given level. Each single class or school did not have to have 8 as average, although it was often 'the understanding' in practice. Usually a university average of more than 9 was considered to be quite good, but what really counted in higher education was the grade obtained in the master's thesis (*Speciale*). There is no system of 'first class degrees', etc. As a classification system, the 13-scale covered the whole educational sector and was unaltered even when its use was changed from relative to absolute.

In addition to problems related to the use of the same scale both for relative and absolute use, further problems emerged. One was that the scale is not compatible with the ECTS scale. One reason was that the expected frequencies of the 13-scale did not fit those of ECTS (Danish Ministry of Education 2004: 52). (See Table 2).

**Table 2 Expected frequencies of the old Danish 13-scale**

| 13 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 03 | 00 |
|---|---|---|---|---|---|---|---|---|---|
| 1% | 5% | 10% | 20% | 27% | 20% | 10% | 5% | 1% | 1% |



This meant that it was not possible to make a unique translation back and from other countries' grading scales. For instance (Danish Ministry of Education 2004), 13 and 11 both translated into an 'A', but the other way around gave rise to problems, since how does one then translate 'A' into *one* specific number at the 13-scale? Another problem was that the 13-scale was used differently in different subjects. For instance, mathematics tended to use the whole scale but not so much the middle grades, whereas, for example, Danish tended to use mainly the middle grades. Also, the various school levels used the grades differently. There was also a grade inflation, for instance, as seen in the increasing use of 13, which is an exception-grade only to be used in very extreme cases of students being very far above what can be expected. In fact, 11 was the actual top grade, which was frequently misunderstood abroad. The 13-scale also tends to be used as an NRA scale even when it is supposed to be a CRA scale. A Grade Commission of the Ministry of Education therefore recommended in 2004 (Danish Ministry of Education 2004) that the 13-scale should be abolished and replaced by a grading scale comparable to ECTS. In 2006, the Ministry of Science, Technology and Innovation sent out a Departmental Order (no. 886) stating that a new grading scale should be used for university examinations from September 2007. The new grading scale was first used in the high schools (*gymnasium)* from August 2006, and at the other levels from August 2007. The new scale is seen in Table 3.

### *Why does the new grading scale look the way it does?*

The Grade Commission made five demands for the new grading scale. First, *internationalisation*. The new scale must be compatible with the ECTS scale to make it possible to translate back and forth between the grades and make it easer for Danish students to be mobile. Second, the new grading system should be used *uniformly throughout the education system* and a student's *grade should express the degree to which the student meets the intended learning outcomes.* This is a change from a content-controlled grading system (and from content-formulated course syllabi) to an objective-controlled system

**Table 3 The new Danish grading scale, the '7-steps-scale' (sometimes called the 12-scale), translated from Danish (Brabrand & Dahl forthcoming)**

| New Grade | Description | ECTS equivalent |
|---|---|---|
| 12 | For an excellent performance which completely meets the course objectives, with no or only a few insignificant weaknesses. | A |
| 10 | For a very good performance which meets the course objectives, with only minor weaknesses. | B |
| 7 | For a good performance which meets the course objectives but also displays some weaknesses. | C |
| 4 | For a fair performance which adequately meets the course objectives but also displays several major weaknesses. | D |
| 02 | For a sufficient performance which barely meets the course objectives. | E |
| 00 | For an insufficient performance which does not meet the course objectives. | Fx |
| – 3 | For a performance which is unacceptable in all respects. | F |

where grades are given in relation to clearly stated Intended Learning Outcomes. It is argued that this will ensure a uniform use of the grades, which is a condition for strengthening collaboration between subjects. Third, the *same scale should be used in the whole education system* regardless of subject matter to enhance the vertical and horizontal mobility between education programmes. This is different from the second point, which mentioned *how* the scale should be used – this third point focuses on *where* the scale should be used. Fourth, there should be a *clear distinction between the steps* but there should still be a high enough number of grades in order for the scale to also be used in formative evaluation. Fifth, it should be possible to *calculate an average*, which means that numbers are necessary and not just letters. Figure 1 illustrates how the Grade Commission came from the ECTS letters' distribution of grades to the numbers in the 7-steps scale.



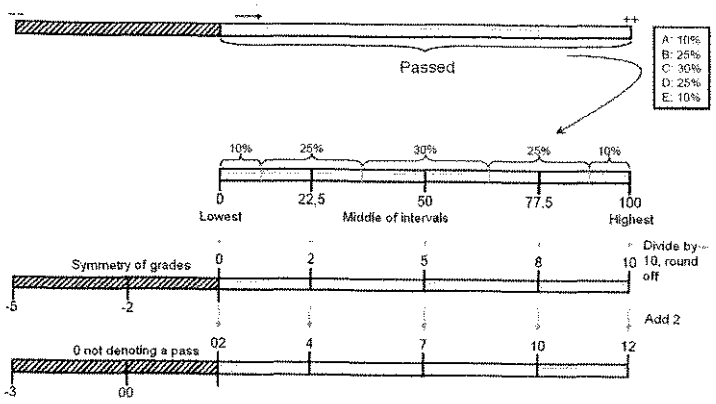


**Figure 1** Illustration of the relation between ECTS and the new Danish scale (Brabrand et al. 2007, based on Danish Ministry of Education 2004: 43–44)]

*Grading traditions*

Grades are always decided between the course teacher and either an internal or external examiner. A university department decides by itself, within some national rules, what courses should have which type of examiners. In 1993, a national corps of external examiners was introduced by a Departmental Order (no. 332). A third of the examiners were to be from outside the university world. A subject's corps of examiners is formally appointed by the Ministry for Science, Technology and Innovation for a 4-year period. In 2006, the Minister of Science suggested that this system should be abolished, partly since it was too expensive, and partly because it was different from what other countries do. But there were many protests from the universities and employers' organisations that did not want to go back to the situation before, where each university had its own corps of examiners, since this might jeopardise a uniform grading.

*The implementation of the new grading scale*

Different universities, and even faculties, implemented the new grading scale differently. For instance, at the Faculty of Science at the University of Århus, the dean determined that new syllabi should be formulated as ILOs, using the SOLO taxonomy relevant to research-

based university teaching and consisting of five levels of progression. First a student becomes capable of dealing with single aspects in a unistructural or multistructural manner. Later the student experiences a qualitative improvement when aspects integrate to form a relational structure; and at the highest level this structure is contextualised and reflected upon (Biggs 2003: 34–53). The use of the SOLO taxonomy had been recommended by the university's higher education network. The Faculty of Science at the University of Southern Denmark had a similar approach (Brabrand & Dahl forthcoming). Other universities took different approaches, some used Bloom's taxonomy (Bloom 1956), which is different from the SOLO taxonomy. Bloom's taxonomy describes the levels of educational objectives moving from lower order to higher order complexity. In the cognitive domain, they are described as knowledge, comprehension, application, analysis, synthesis and evaluation. Other universities leave it up to individual departments or faculties to decide on how to formulate the ILOs.

*Experiences*

The use of the 7-steps scale is very new in higher education, but the scale has been used for only a year in high school where it has met a lot of criticism. One of the criticisms (Wissing 2006) is that the scale is not good for formative evaluation due to the fewer number of grades to choose from, the large steps between the grades – particularly from '7' to the neighbouring grades – and the presence of a negative grade. Also in relation to '7', it used to denote 'below middle' in the 13-scale whereas in the 7-steps scale it is a better grade; in fact, official grade converters by the Ministry of Education state that a '7' in the new scale is equivalent to a 8/9 in the old scale – hence now '7' is a kind of 'above middle'. But students are used to '7' being not so good. It has also been criticised that it does not fit today's society's demand for creativity and invention to have learning objectives clearly predefined and specific and a grading procedure very fixed on what is 'missing' from the predefined descriptions of the goals. The examiners also run the risk of being very focused on what is 'missing' due to the explicit use of these expression in the explanation of grades. Also some studies



at universities miss the 13-grade in order to really find and award the really able students. However, others also state that the scale is easy to use, one quickly finds the appropriate grade since there are fewer grades to choose from.

## Norway

The ministry of higher education has given each institution a right to award a specified set of degrees. Included in this right, the institution has its own legal basis for conducting the examinations. Within this frame, the university faculties do both the formal and practical work of arranging examinations, making the assessments and setting the grades. The faculties themselves decide whom they will use as examiners. Up to 2003, almost all examinations had their assessments done by an internal examiner (the person with academic responsibility for the course) together with an external examiner. Since 2003, the use of external examiners was no longer required, except for the master's thesis. Hence, the use has been reduced, mostly for economic reasons. External examiners are now used mostly in the preparation of the examination questions.

The upper secondary school in Norway has for the last 40 years used a grading system based on numbers: 6 as the highest ranked, 1 (or even 0) as not passed. This system is still in use, and has not been influenced by the reform of grading in higher education.

*Previous grading scales*

Up to 2003, universities and university colleges in Norway usually adopted a grading scale founded in the national academic tradition: a number-based scale together with a Latin term. The scale had 1 as its highest mark and 4 as the lowest pass mark. Each integer step was mostly divided into 0.1 size steps, making it a scale with 31 pass marks! Fail was termed *immaturus* or was marked with a number above 4, up to 6. This meant that even failed results could be ranked. During the last century, new academic institutions and subject areas have grown and started their own grading traditions.

The usage was even more complex, apparently identical scales were used in different ways – depending on the tradition in different academic groups. You had to know the specific tradition in the subject area to interpret the value of a mark. For instance, a 2.2 mark in the humanities and social sciences was a really good result, but in the natural sciences the same mark would be satisfactory, but not outstanding. Common for all scales is a large amount of steps. The scale most used had 31 different steps for passed results (1.0–1.1 - ... - 3.9–4.0). The 1.0 mark was very seldom used in the humanities. When closing down the scale in 2003, academics in the humanities were more or less still waiting for the ultimate genius. In natural science and technology, the 1.0 mark was far away from being a rare mark.

Up to 2003, each institution decided which grading scales to use. The decision was usually linked to traditions. In 2002, the Ministry decided, based on the new law for higher education, that all institutions should choose one of two grading scales: either a 6-step scale A–F or a 2-step pass/fail scale. This was realised for all not later than the autumn semester of 2003.

### *Political reasons for creating a new grading scale*

The need for communicating the candidates' ability to employers was the prime political reason for revising the variation in grading scales. This variation was annoying, for students as well as for employers. Students no longer follow pre-decided pathways to professions on the basis of their degree programme, and there has been a growing tendency to switch between different positions in a more flexible labour market. Students also have a tendency to mix subjects from different institutions in higher education that have different grading traditions. But probably the most important reason for a revision was that Norwegian higher education institutions needed a simpler and more transparent grading system because of their expanding contact with institutions abroad.

In 1999, the four universities in Norway (at that time) established a committee that was asked to draw up a proposal about how a new, common grading scale could be designed and described. The





committee had members representing both scientific and administrative staff. Since the ECTS scale was already established, though it was primarily for transferring grades from one national system to another, the committee saw that this scale could be a good basis for the new common national grading scale. The ECTS system had a reasonable number of grading steps, a reasonable use of symbols and was well known in Europe. The committee stressed very clearly that the national scale should be an absolute (CRA) scale, with qualitative descriptions, and was not to be mixed up with the relative (NRA) ECTS scale, which was made purely for the purpose of translating the original grades between countries, based on a preset quantitative grade distribution.

The proposal from the grading committee (Norwegian Council of Universities 2000) was sent to the official committee considering the educational structure of higher education in Norway, which at that time had reached the concluding stage of their work. The grading proposal was included and was approved by parliament as a part of the Act relating to higher education in Norway in 2002.

### *The new grading scale*

At the start of the academic year 2003/4, if they had not already done so, all the institutions of higher education in Norway were required by law to convert to the new common grading scale. Table 4 shows the qualitative description given for each grade.

This is the common national and general description, used for all subject areas. For each discipline a specific description is made, stating what academic criteria in that subject the candidate should fulfil to be awarded grade A, B etc. This level of description is based on a common national agreement.

As a tool for quality assurance, the *quantitative* distribution shown in Table 5 has been set.

This could indicate that the new scale is a relative (NRA) scale. But the scale used for direct assessment is the one as described in Table 4. The intention for the percentage distribution as in is that, after some time in

**Table 4 The Norwegian grading scale since 2003**

| Symbol | Description | General, qualitative description of valuation criteria |
|---|---|---|
| A | Excellent | An excellent performance, clearly outstanding. The candidate demonstrates excellent judgement and a high degree of independent thinking. |
| B | Very good | A very good performance. The candidate demonstrates sound judgement and a very good degree of independent thinking. |
| C | Good | A good performance in most areas. The candidate demonstrates a reasonable degree of judgement and independent thinking in the most important areas. |
| D | Satisfactory | A satisfactory performance, but with significant shortcomings. The candidate demonstrates a limited degree of judgement and independent thinking. |
| E | Sufficient | A performance that meets the minimum criteria, but no more. The candidate demonstrates a very limited degree of judgement and independent thinking. |
| F | Fail | A performance that does not meet the minimum academic criteria. The candidate demonstrates an absence of both judgement and independent thinking. |

**Table 5 Distribution of Norwegian grades**

| Symbol | Description | ECTS distribution | Frame values | Ranking criteria |
|---|---|---|---|---|
| A | Excellent | 10% | 8–12% | Clearly above average |
| B | Very good | 25% | 20–30% | Above average |
| C | Good | 30% | 24–36% | Average |
| D | Satisfactory | 25% | 20–30% | Below average |
| E | Sufficient | 10% | 8–12% | Far below average |

use (a minimum of 3 years), a complete record of programmes of study in a particular subject area covering a great number of candidates should end up with a distribution of pass grades within these frame values. If so, the grading scale is being used as intended. If not, it is probably an indication that the qualitative description for this discipline should be revised. In a letter dated May 2004 to all higher education institutions in Norway, the ministry clearly stressed that if the scale is used as intended, there would be no need to translate Norwegian grades into ECTS grades when comparing with other national scales. This letter is the only example of ministry involvement after the law was approved.





*How the new grading scale has been implemented*

The Norwegian Council for Higher Education, UHR – a voluntary co-operation among higher education institutions founded in 2000 – has co-ordinated the implementation. The ministry seemed to be very satisfied with this initiative. This might be interpreted either as the ministry was hoping that the institutions themselves would solve any conflicts and problems or as a vote of confidence in the professionalism of the universities and in favour of decentralisation. UHR's follow-up was to establish several national panels, covering different disciplines. The mission was to review the use of the scale for each discipline by making statistics and analysing the grade distribution, to see if the scale was used as intended. All institutions deliver their own examination results, which are gathered into a national statistical database.

*Experiences*

After five years of use, the new common grading scale has become the national scale, as intended. Students and professors – as well as institutions outside higher education – have become familiar with it. The transition did not result in the turbulence that some expected. However, some problem areas have been uncovered: First, using As and Bs is confusing for some of those who are in contact with higher education in the United States, where they have quite another distribution and use of these grades. However, this is quite a limited problem that is mostly solved by a diploma supplement. Second, it seems that examinations in higher degree programmes have a grade distribution that covers more of the As and Bs than lower degree examinations. The discussion has therefore been: should the description

take into consideration the fact that higher degree students are more experienced than lower degree students? Or should all examinations within a discipline be looked upon as a unit – giving credit to those who have been studying for a longer period?

In order to give an answer to this, the UHR in 2005 established a specialist group for collecting and analysing the statistical data from all the institutions, in order to give a picture of how close the usage is to the expected distribution in Table 5. The analysing group made reports for each year. This initiative was followed with great interest by the higher education institutions. The conclusion so far (autumn 2008) is that when statistical data for all institutions are sampled, the total distribution was quite close to what was expected. The grade distribution most close to the ideal distribution is for lower degree examinations. Higher degree examinations in total intended to have a higher amount of As and Bs – but not in an alarming way. But picking out the master's theses, the picture was quite different: nearly 80% of all theses were assessed as A or B. To follow up these statistical results, the UHR in spring 2008 established a new national committee. During autumn 2008 this committee is due to analyse the statistics of grade distributions for the first three years, to see if there is a need to revise the set of general qualitative descriptions. And as a special task, give advice on what to do with the 'misdistribution' of the grading of the master's theses. But it is interesting to observe that this was not stressed as a problem when the old grading scales were used. Probably, we would argue, converting to a new scale has uncovered latent conflicts or perhaps created new ones when the distribution is expected to fit an NRA perspective.

## Sweden

*Swedish schools – from NRA to CRA and a proposal for 'ECTS-like' grading*

In Sweden, grading in schools and in higher education are treated as different policy domains. Grading based on Norm-Referenced Assessment was applied both at the compulsory school level and in upper secondary education from the end of the 1960s until 1994, when a shift to Criterion-Referenced Assessment took place. In the 1960s, the introduction of NRA, with a grading scale of 1–5, having 5 as the highest grade, was underpinned by extensive educational research and was part of the intention to create equal opportunities in the development of a welfare society. The long period of grading based in NRA also meant that much of Swedish educational debate and research bypassed the movement to frame educational activities by behavioural goals or student learning objectives. Since the 1970s, Sweden developed a quite grading-hesitant educational culture. Children at compulsory school level are not graded at all until they reach their eighth year in school. The shift from NRA to CRA in schools, which also included a shift to grading scales with only three or four levels, turned out to be a slow and complicated process (Selghed 2004). In March 2007, the new right-wing/liberal government declared a marked shift in the old policies by setting up a committee to introduce both earlier grading at the compulsory level and a new grading scale with seven levels 'like the ECTS' in schools. Two main arguments were put forward: That an increased number of steps would support 'better and more rewarding learning' for the pupils and that parents would get more detailed information about the results of their children. The proposal was presented in February 2008 (Ministry of Education and Research in Sweden 2008) and outlined a grading scale in six steps, A–F, where F is failed (The Fx in the ECTS grading scale was rejected, thus six steps). When expressed in numerical values, E (pass) should be worth 10 marks and A, 20 marks, and each level in between equally added 2.5 marks. The government bill was presented in November 2008. If the scale is approved, it will be launched in 2011. However, this was only aimed at compulsory education and upper secondary levels. The story of grading in higher education looks quite different and is not easy to compare with the development in Denmark and Norway either.

*Higher education in Sweden*

Towards the end of the 1960s, Sweden created a unitary system for higher education, which besides traditional university education also

included most vocational programmes, such as nursing, paramedic education and teacher training. A general credit system was introduced, which led to a rapid shift to modular structures. The whole sector is covered by a comprehensive legal framework in the Higher Education Act and Ordinance. About 95% of higher education and research is carried out at the 14 state universities and at the 22 state university colleges. There is also a small group of private higher education institutions partly funded by the state (Ministry of Education, Research and Culture in Sweden 2005 & 2006).

The explicit political intention, supported by a long series of social-democratic governments, was the endeavour to support unity and uniformity (*enhetlighet*) across all institutions in Sweden. An example of this policy is the centralised system for admission to programmes at the national level, with common entrance requirements across all institutions for most programmes. From 1977 to 1993, the structure of programmes in higher education also was firmly regulated at the national level by centrally decided 'study plans'. The number of students admitted into different programmes at each institution was decided by national authorities. Thus, the control of the system was mainly exerted by the input factor: *planning* based in the common structure. In 1993, a major deregulation was launched. The responsibility for programme structure was transferred to the institutions. An impressive range of new degrees and courses was created after the deregulation of the system (Bauer, Askling, Gerard Marton and Marton 1999: 258).

### *The examination system and grading traditions*

The 'course' became the basic entity of all programmes in the 1960s, and still keeps this status. A course consists of a module with a syllabus. *Only* courses are assessed and graded. The single teacher is the 'examiner' and has the full responsibility for the assessment and grading of students at the end of the course. Thus, in Sweden, all assessment is framed by a course, and assessment practices are integrated into the teaching and learning of each course/module. No concluding or final exams are allowed for general or professional degrees. There are no traditions of external examiners at degree level. The decision to award a full degree is therefore basically an administrative one, verifying that all the required courses are registered as passed. Grades from different courses are not aggregated or summarised to a final grade, which means that the student gets no final mark or grade linked to their diploma (Lindberg-Sand 2005).

From 1993 onwards, the formal curricular system in Swedish higher education thus can be described as a decentralised, loosely-coupled and modularised system, with very few national regulations or policies concerning teaching, assessment or exams. In most higher education institutions, the decisions concerning syllabi and assessment are made close to the teachers at the department or unit level. The system relies on extensive teacher responsibility for assessment and grading and for the design and maintenance of syllabi. Very few institutions have developed policies of their own concerning examination and assessment. These aspects are mostly handled at faculty, or even at department level.

The nature of the Swedish grading system is not formally defined at the national level as either Norm-Referenced Assessment or Criterion-Referenced Assessment. The scale considered normal is fail/pass/high pass, but higher education institutions can choose any other grading scale or grading system they want. From 1977, higher education institutions could also choose a grading scale with only pass/fail if they so wished. Several important programmes leading to recognised professions such as doctor, psychologist or social worker, grade all their courses as only pass/fail. Thus, large groups of academic teachers have no professional experience of designing and marking assessments with grades other than pass/fail. Grading practices vary within institutions, since faculties and departments can have their own system, but grading of credits more specific than 'high pass' is uncommon. The different traditions were, until 2006, mostly linked to the field of study. For instance, most programmes in Engineering used a grading 1–5 (where 1 and 2 were failed), whereas Faculties of Law mostly

used the old grading-scale Ab-ba-B-C (where C was failed). The varied traditions also meant that large universities, with a great variety of study fields and programmes, normally did not interfere with faculties or departments by trying to compare or synchronise grading processes or assessment traditions from different fields of study. In this respect, the situation might look similar to that in Norway before their reform. But there are two major differences: First, the simple Swedish grading scales had not created the same problems with internal comparisons as the complicated Norwegian ones. Second, the legal construction of assessment in Sweden defines examination as a part of each course only. There is no tradition of or organisation for handling examination at the level of programme or degree. The distribution of grades is not followed up at national or institutional level and is not a part of the system for national quality assurance.

*The Bologna process and the Swedish grading system*

When the first Government Report concerning the Bologna Process was presented (Ministry of Education, Research and Culture in Sweden 2004), it put forward that explicit Intended Learning Outcomes for all degrees and for all courses should be developed and a new credit system introduced. Furthermore, it suggested that the various grading scales used in different fields of study should be supplemented with 'ECTS grades' with seven steps, but based on Criterion-Referenced Assessment. Any student could then ask to have such an ECTS grade as well as the original grade. The main argument was that these changes would support the international recognition of Swedish higher education and facilitate the mobility of Swedish students. The proposal raised a fierce debate along two lines, since the report labelled the suggested CRA grading scale as 'ECTS grades'. Many thought this was strange and quite improper, as the ECTS scale is based on Norm-Referenced Assessment and is only intended to be an explanation of the original grade. If carried through, the proposal might force institutions to work simultaneously with two original CRA grading scales for the same group of students. The other part of the debate concerned the



Swedish assessment culture, and consisted of clashes between negative and positive conceptions of the grading of students at several levels (Lindberg-Sand 2007). Since all the old grading scales had fewer steps than seven, many voices in the pedagogical debate expressed doubts about the value of increasing the number of levels. They highlighted the possible detrimental effects on student learning when teachers are forced to change their assessment practices to put the marking of assessments at the centre of attention instead of designing assessments for learning (Dahlgren & Fejes 2004; Silén & Harrysson 2004; Lindberg-Sand 2005). The resistance was especially outspoken in medicine, where the ongoing development of problem-based learning was nurtured by the simple pass/fail grading. But when at last, in 2005, the Government bill (Swedish Government 2005) was presented, it did not put forward any suggestions concerning grading at all. When all decisions were finally made, in the end the national legislation connected to grading in higher education was not changed in any way.

*Results of the reform: grading is changing – but how?*

When the reform was launched in 2007, Swedish assessment traditions and grading were initially at the centre of attention. When the Government bill did not include a new grading system, these discussions ended or were replaced by all the other aspects of the rapid implementation of the reform. But a very noticeable outcome of the reform is a new tendency for institutions to decide on the new grading systems that are to be applied to all of their programmes – thus an institutional centralisation of grading scales. For instance, Stockholm University, together with some other institutions, has decided to implement a CRA grading scale with seven steps. At the same time, Uppsala University has chosen a grading scale with four steps based in CRA, while several other institutions have decided on three levels, and yet others, like Lund University, have kept the tradition with different scales in different fields of study. But here too some fields of study have chosen to change to a seven-step scale or to give all of their students an ECTS grade based on NRA. The variation has

increased and comprises new paradoxes, since the same professional degree studied at different institutions can now have a grading in two, three, four or seven levels. There is also a variation concerning how institutional grading scales are linked to CRA and NRA, since it is optional for institutions to choose which system to use. Probably the topic of grading will reappear on the agenda when the students start to compare their situations.

## Discussion and conclusions

In this section we will focus on four issues that emerge when comparing the three countries, namely, the different grading systems, the level of state control, the difference between NRA and CRA types of assessment related also to the ECTS and finally, the question of what 'ECTS-like grading scales' are exactly.

### *Different grading systems*

Table 6 illustrates the different grades used before and after the change of grading system in the three countries.

It is seen that although the situation in the three countries was different 'before', it is by no means the same 'after'. In Denmark and Norway, one of the main purposes of having an ECTS-like scale was

**Table 6 Overview of the grading scales in Denmark, Norway, and Sweden**

| | Denmark | Norway | Sweden |
|---|---|---|---|
| Grades 'before' | Same scale used differently – more than seven levels | Many different scales – more than seven levels | Many different scales – two to four levels |
| Grades 'now' | ECTS-like scale, with numbers | ECTS-like scale | More diversity – some institutions have ECTS-like scales |
| Direction of change: number of levels | Decreasing | Decreasing | Some – No changes<br>Some – Increasing |



to make it easier to compare grades between countries. Both countries therefore seem to have 'merged' into the ECTS ideal, but in different ways. They now have grading scales that seem alike, except that Denmark uses numbers calculated inventively from the anticipated distribution of the translation to ECTS grades (see Figure 1). However, while earlier the grading systems were obviously different, easily seen in the variety of scales and letters, they are today only seemingly alike. The common use of the ECTS translation tool and reference to ECTS language disguises the fact that behind the ECTS letters is a different interpretation of the grades. First, the Norwegian qualitative description of evaluation criteria is not formulated in relation to the intended learning outcomes of the course or unit, as recommended by the Bologna Process, while the Danish grade explanation is. The Norwegian descriptions were made before the learning outcomes in the Framework for Qualifications in the European Higher Education Area (EQF-HE) or the European Qualifications Framework (EQF) were agreed upon, and have not yet been revised to consider these. The different descriptions also indicate quite different educational and philosophical approaches to grading (and teaching). Hence, for instance, a Norwegian 'A' does not reward the same things as a Danish 'A' – ergo an 'A' is not always an 'A'. The ECTS-like grades are therefore examples of 'boundary objects' (Bowker & Star 1999). When a Danish 'A' is compared to a Norwegian 'A' in a certain community of practice, they may well be treated as if they are exactly the same, although they are produced in two very different classification systems. They share only the label, but in this community it may be only the label that is requested for the practice to fulfil its mission.

Second, the different use of numbers linked to the ECTS letters will, in some cases, result in very different averages, as seen in the example below. In Table 7, we see the invented students, D (Danish) and N (Norwegian), who have both studied the same topic and received the same four grades in courses that carried the same number of ECTS credits.

The reason for the different averages is that the Norwegian scale has similar numerical distance between the letter grades, which the

**Table 7 Example of same national grades giving different ECTS-average**

| | Student D | Student N |
|---|---|---|
| Course Grades | 12 12 4 4[#] | A A D D[##] |
| Translated into numbers | | 5 5 2 2 |
| Average | 32/4 = 8 | 14/4 = 3.5 |
| Translated/rounded off as an ECTS letter | C (7) | B (4) |

[#]Denmark: (12, 10, 7, 4, 2, 0, −3) = (A, B, C, D, E, Fx, F)

[##]Norway: (A, B, C, D, E) = (5, 4, 3, 2, 1)

Danish scale does not. When *numbers* are used, a comparison of the scales reveals that they are in fact different classification systems (Bowker & Star 1999). In fact, the proposed Swedish system, which in the future may also be used in higher education, operates with yet another system of numbers (20; 17.5; 15; 12.5; 10) = ( A, B, C, D, E) for the passing grades. However, since the numerical steps between the grades are of equal size, this proposed system is similar to the Norwegian system; hence, a Swedish student would also get B for the same grades. Since averages could be calculated, it turns out that the numbers assigned to the grades are far more important than the letters A–F. As long as the students are in their own country, this does not cause problems, since people are familiar with the specific national grading system and the same system has been used for all students. However, a main argument for the Danish scale change was to secure Danish students' easier access to foreign universities than before, where '13' was often misinterpreted as 'A'. But the two ECTS-like scales in Denmark and Norway are not alike, which means that in competition with Norwegian students, unfair comparisons might still occur. What further complicates the situation is that in Sweden, the institutions that have changed to ECTS-like scales may have designed their scales differently, since they can choose to base the grades on NRA or in CRA or in combinations of both, and also decide for themselves how to number the spaces between steps. But since the task of calculating average grades is actually not applicable in the Swedish grading system, this might not be recognised at all. The grade for



each course is a final whole and there are no legal means to calculate the average grade for a cluster of courses or for a degree. Each of the Swedish course grades has to be supplemented by a separate ECTS grade. Hence, the new ECTS-like grading scales have not solved the problems of transferability between grades, or even that of a stable translation of grades.

### *Level of state control*

As described above and set out in Table 8, the path to new grading scales was different in each of the three countries.

We see that Denmark has a high level of state control, where the state initiated the change of grading scale by law even though it let it be up to the single universities to implement it. The decision, however, covered the whole education sector. Sweden seems to be at the opposite end, with a low level of centralised state control, when looking at assessment and examination in higher education. The Ministry's failed attempt to include a new grading scale in the Government Bill kept the initiative clearly at the institutional level. A voluntary collaboration between Swedish higher education institutions is organised in SUHF (The Association of Swedish Higher Education). A proposal for institutional collaboration with regard to a CRA-based grading scale in four levels was rejected by SUHF. A recent search on the official parts of the SUHF homepage ( < http://www.suhf.se/ > ) gives zero hits for 'grades' or 'grading scales'. In Norway the initia-

**Table 8 Overview of the different steering instrument and scope of the decisions in Denmark, Norway, and Sweden**

| | Denmark | Norway | Sweden |
|---|---|---|---|
| Source of Initiative | Government | Institutional collaboration between universities | Higher Education Institutions |
| Educational sector covered by the new scale | Education as a whole | Higher Education | Higher Education Institutions or faculties |

tive was born and bred by the collaboration between four Norwegian universities and the government implemented a suggestion from this group to adopt a scale based on the ECTS scale. The overall tendency, however, is that of an increasing centralisation, where new grading scales are covering larger educational domains. This is seen both in Norway and Sweden.

Even with a high level of state control, as in the Danish case, the implementation is still decentralised. This means that in practice even though Denmark has a centralised system of external examiners, an 'A/12' at one university might not be equivalent to an 'A/12' at another university – even if the programme and course title/theme is the same, and at the same stage (such as a graduate or an undergraduate course). This is because grades are now to be given based on the extent to which a student meets specified Intended Learning Outcomes; but the ILOs have been formulated differently at different universities and faculties, using different taxonomies or no taxonomy at all. As stated by Dahl (2008), even if every single university course syllabus was formulated as ILOs using the same taxonomy, different universities have different teaching styles, foci and assessment practices. Some have a large degree of problem-based learning in groups, others are traditional, and others are in between, so that even courses that are named the same are different. Hence, similar grades do not necessarily correspond to similar learning outcomes. We would also argue that the kind of coordination made possible by applying the new frameworks for qualifications mentioned earlier (EQF-HE & EQF) must be characterised by a mutual interpretative translational process built on trust and not on identical syllabi (Lindberg-Sand 2008). When the focus is on separate grades, to what extent the grade can be technically translated from one nation to another becomes rather irrelevant. Such a translation cannot ensure a transparent and fair situation.

### *Norm-Referenced or Criterion-Referenced scales – classification, control and quality*

The ECTS grading scale is presented as an NRA scale, but it is only meant to be a supplementary explanation to the student's original national grade. The 'ECTS grade' is thus added to the original grade.



What an ECTS grade may display must therefore be judged by looking at the original grade and the supplement together as *a pair*. When the ECTS-like scale is adopted as the original national grading scale, it becomes unclear what this label 'ECTS' means, especially when the national scales are also declared to be based in CRA. As our example above shows, there are several options. Both in Norway and Denmark, the ECTS-like scale is used as a CRA scale. This might seem as a contradiction, but the way both countries have chosen to deal with this is to emphasise that the distribution 10–25–30–25–10% is something that will hopefully be seen over a number of years, it is *not* something that should be of concern for examiners when they give grades. In other words, the purpose is not to see this distribution within each single course, programme or even year; instead, it is expected that this pattern will be revealed over the years. Still, to make the normal distribution of grades over the total population of students an inbuilt aspect of the grading system also supports a certain view of learning that includes restrictions on the possibility to succeed. And, of course, academics are aware of how the grading system is supposed to work and may adapt their assessment to the perspective on human learning underpinning the distribution. From this perspective, 10% of all students in master's programmes will have to be relative failures (grade E) in the Norwegian grading system – a point that some academics have questioned.

Both in Denmark and Norway, it is stated that if this pattern does not emerge over the years, it will be necessary to adjust the (absolute) criteria for getting each grade so that this pattern will appear later. This is to state that the ideal distribution of grades is a more fundamental aspect of the classification system than the design of the intended learning outcomes and the criteria in relation to the intentions and conditions for student learning. Time will tell whether this pattern will indeed emerge, and in case it does not, how Denmark and Norway will act – and what the action will be. In both Denmark and Norway this means that nationally or institutionally based quality systems will have to include statistical reporting of distributions of grades, so that they can be compared. This question is not even

discussed or regarded in the Swedish decentralised grading system or in the ongoing changes of the national quality assurance system (Swedish National Agency for Higher Education 2007). One can say that in Sweden the topic of grade distribution is invisible at most educational levels and also, regarding the fields of study where only pass/fail grading is practised, is a topic totally void of valuable information.

The mixing of NRA and CRA perspectives on learning underpinning different grading systems, creates some delicate problems. If, for instance, groups of students in a certain field of study get better grades than before, this obviously can have two explanations. From an NRA perspective, where the normal distribution is a goal in itself, this may be apprehended as a lack of quality indicating ongoing grade inflation. From a CRA perspective it will be judged as if the quality is increasing: students have achieved the learning outcomes to a higher degree. But the grading scale in itself can never guarantee the correctness of any of these interpretations. And furthermore, if teachers in Denmark or Norway are implementing Criterion-Referenced Assessment for grading as the new scales require, and if the distribution of grades turns out not to follow the 10–25–30–25–10% pattern, the universities would either have to raise or lower standards in order for this pattern to emerge. Hence, the comparison of students from even the same university programme *across the years* becomes rather impossible. For instance, a 'B' student one year might be as knowledgeable as an 'A' student the next year. When single grades become more valuable for the individual, and in accumulated forms also for the institution, this calls for increased efforts to develop methods for quality assurance and educational research of student assessment in higher education, which do not take the intentions behind the actual grading system for granted.

### *What are 'ECTS-like grading scales' and how do they relate to ECTS grades?*

The three Scandinavian countries in this study have all relied on, and replied to, the supplemental grading scale that is a part of the ECTS,



in order to shape their new policies concerning grading in connection to the Bologna Process. Looking at the transition of the ECTS grading scale to a national or institutional grading scale, several differences and also difficulties have been displayed, when the grading scales are regarded as classification systems with multiple functions at several levels of the educational system (Bowker & Star 1999). Some of the difficulties are related to the relationship between the grading scales and NRA and CRA. The specific design of the grading scale, when levels are named or numbered, may also result in a different classification of the learning outcomes evaluated in assessments. This study suggests that the only common trait signifying all ECTS-like scales is that the scales contain five levels for grades passed, and nothing more. Therefore one cannot really say that a country 'has an ECTS grading system'. At most, it has a model of five passing grades and one or two failed grades. As seen in our three cases, these can be constructed in many different ways; so the need to add the supplementary ECTS grade to explain the original grade still exists. And one should also note that if the ECTS-like original grade is based on CRA, this will mean that a correct application of the ECTS, based on an NRA distribution, might result in a very different ECTS grade being added to the original ECTS-like grade! In fact, the Danish diploma adds the 'corresponding' ECTS grade next to the number grade from the new national scale. For instance, '7' is always followed by a 'C'. However, this is not in line with the relative construction of ECTS unless, in fact, one later learns that the grades were in fact distributed exactly as 10–25–30–25–10%. If the new grades do not follow such a pattern, a '7' might just as well be translated into a 'B' (if 'not enough' got the higher grades '10' and '12') or a 'D' (if 'too many' got these higher grades).

One thing that we have not touched upon yet is in what way the ECTS grade is constructed in relation to the original grade. This turns out to be an intricate matter. One could say that one assumption behind the supplementary and explanatory ECTS grade is that whoever assigns this added feature to the original grade has to have access to reliable information about the rank order of the student in relation to the course grades as a whole, even so in a CRA-

based grading system. Otherwise the intention of the ECTS grade to explain the location of the original grade within the distribution of all grades for the course will be thwarted. This kind of information is not always available, especially when CRA-based grading scales with fewer than five passed steps are used. The situation in Sweden can be used as an illustration. In the absence of a national decision regarding grading scales, some institutions are introducing a seven-step ECTS-like grading scale as their original one, as shown in our cases. Other institutions have kept their three- or four-step original grading scales and together with this original grade have also given all or some of their students an ECTS grade in the seven-step model. But this is done in different ways. Some institutions have decided on an administrative procedure where the original grade automatically is translated to an ECTS grade, though they do not control the distribution. Others have taken another route. Since the information contained in the original grade is not sufficient to add the ECTS grade (since this grading scale has more steps), the assessment procedures are also adapted to the assignment of the ECTS grade. One could assume that departments or faculties that are working for a change in the Swedish grading tradition have utilised this as a strategy to introduce grading with a larger number of steps than their institution has formally decided upon. By doing so, the ECTS grade is no longer only an explanation of the original grade, but is becoming an original grade in itself. Presently this causes quite a lot of confusion among Swedish higher education institutions, since no one really knows what is being referred to when talking about 'ECTS grades'.

When the curricular aspects of the Bologna Process and the new frameworks for qualifications (EQF-HE & EQF) are implemented in national frameworks, we might perhaps assume that in Europe grading based on CRA will also increase. If so, the problems with the relationship between the original grades and the ECTS grades described above will persist, especially if nations are converting to simpler scales with almost the same number of steps as the ECTS scale. Then one might ask if the grading tool in the ECTS will have fulfilled its initial mission and can be abolished.



## Final comments

In this paper we have described and discussed how Denmark, Norway and Sweden each have responded to the challenge of the Bologna Process to change their grading scales in higher education. Each country largely referred to internationalisation as the political reasons for the change of scales. We saw that even though the three countries are quite similar in many areas, from the outset they had very different assessment traditions and grading scales, with Denmark being much centralised, Sweden much decentralised and Norway in the middle. The new grading scales were all based on the translation tool 'ECTS grades'. And since this tool is based on Norm-Referenced Assessment, all three countries had to struggle with the problem of combining this trait with the development of their new scales, each of which should be based in Criterion-Referenced Assessment. Denmark and Norway made the choice to align the total distribution of grades to the distribution stated for the ECTS grades. Sweden failed in this attempt partly due to the decentralised assessment system, where each higher education institution decides on their grading system. The idea of combining Criterion-Referenced and Norm-Referenced assessment does not work as an insurance against either grade inflation or falling (or raising) standards. Although the three countries have chosen the same source of inspiration for their construction, the actual results – the ECTS-like grading scales – are different when analysed as classification systems. In fact, one could argue that before everybody knew that the scales were different – hence work was required to understand the foreign diploma – now, the scales are seemingly alike and the differences are hidden for most people. Hence, the intention in the three countries to ease the international recognition of the national grades has not yet succeeded, if the condition for the recognition is a real comparability between grades. But if international recognition works on the superficial level, where grades just look the same, mobility may nevertheless become easier. Our work shows that the vision of simple transparency turns out to be an illusion.

**Bettina Dahl** is an Associate Professor in mathematics and higher education at the Department of Science Studies, Faculty of Science, at the University of Århus in Denmark. She holds a Ph.D. in mathematics education from Roskilde University, Denmark.
*Contact*: University of Århus, Department of Science Studies, C. F. Møllers Allé Building 1110, DK-8000 Aarhus C, Denmark. E-mail: bdahls@ivs.au.dk

**Eirik Lien** is Senior Advisor at the Norwegian University of Science and Technology in Trondheim, Norway.
*Contact*: Norwegian University of Science and Technology, NO-7491 Trondheim, Norway. E-mail: eirik.lien@ntnu.no

**Åsa Lindberg-Sand** holds a Ph.D. in education from Lund University, Sweden. She is a senior lecturer in higher education and educational developer at the Centre for Educational Development at Lund University, Sweden.
*Contact*: Lund University, Centre for Educational Development, P.O. Box 118, SE-221 00 Lund, Sweden. E-mail: Asa.Lindberg-Sand@pedagog.lu.se

## Notes

1. We would like to note that the order of authors is alphabetical. All three authors contributed equally to this paper.

2. In a European context, 'liberal' usually denotes 'right wing' economic liberalism whereas in a US context, 'liberal' is 'left wing'. In this paper, 'liberal' is used as in the European context.

## References

Bauer, M., Askling, B., Gerard Marton, S. and Marton, F. (1999) *Transforming Universities: Changing Patterns of Governance, Structure and Learning in Swedish Higher Education*, London: Jessica Kingsley Publishers.

Biggs, J. B. (2003) *Teaching for Quality Learning at University*, Maidenhead: Open University Press.



Bloom, B. S. (ed.) (1956) *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain*, New York: David McKay.

Bologna Process Stocktaking Report (2007) Report from a working group appointed by the Bologna Follow-up Group to the Ministerial Conference in London, May 2007.

Bowker, G. C. and Star, S. L. (1999) *Sorting Things Out. Classification and Its Consequences*, Cambridge Massachusetts: The MIT Press.

Brabrand, C., Sørensen, H. K., Dahl, B., Kjeldsen, H. and Madsen, T. V. (2007) *Den nye karakterskala: Implementation på NAT/AU*, presentation at the Faculty of Science, University of Aarhus, in relation to the implementation of the new Danish grading scale, 24 January.

Brabrand, C. and Dahl, B. (forthcoming). 'Using the SOLO Taxonomy to analyze competence progression of university science curricula', *Higher Education*.

Bryson, L. (1992) *Welfare and the State*, London: Macmillan.

BWG-QF (2005) *A Framework for Qualifications in the European Higher Education Area*, Bologna Working Group for Qualification Frameworks, Copenhagen: The Ministry of Science, < http://www.vtu.dk > .

Dahl, B. (2003) 'Tensions between the European and the Nordic dimension in education, with particular reference to Sweden', in D. Phillips & H. Ertl (eds) *Implementing European Union Education and Training Policy – A Comparative Study of Issues in Four Member States*, Dordrecht: Kluwer Academic, 87–115.

—— (2008) 'International mobilitet med den nye danske karakterskala?' *MONA (Mathematics and Science Education)* 4, no. 2: 110–117.

Dahlgren, L. O. and Fejes, A. (2004) 'ECTS-skalan – att mäta eller mota lärande?', paper presented to the Quality in Swedish Higher Education conference, Linköpings universitet.

Danish Ministry of Education (2004) *Betænkning om indførelse af en ny karakterskala til erstatning af 13-skalaen*, Betænkning nr. 1453, Karakterkommissionen (nov. 2004), < http://pub.uvm.dk/2004/karakterer/ > (accessed 31 June 2008).

Danish Ministry of Science, Technology and Innovation (2005) *Udviklingen på de danske universiteter: Redegørelse til Folketinge, April 2005*, Copenhagen.

Danish Ministry of Science, Technology and Innovation (2008) *Universiteternes nye danmarkskort*, < http://videnskabsministeriet.dk/site/forside/nyheder/Aktuelle-temaer/nyt-danmarkskort-for-universiteterne/universiteternes-nye-danmarkskort > (accessed 31 June 2008).

ENQA (2005) *Standards and Guidelines for Quality Assurance in the European Higher Education Area*, Helsinki: European Association for Quality Assurance in Higher Education.

EU (European Union) (2008a) *The Bologna Process: Towards the European Higher Education Area*, < http://ec.europa.eu/education/policies/educ/bologna/bologna_en.html > (accessed 31 June 2008).

EU (European Union) (2008b) *ECTS – European Credit Transfer and Accumulation System*, < http://ec.europa.eu/education/programmes/socrates/ects/index_en.html > (accessed 31 June 2008).

Fägerlind, I. and Strömqvist, G. (ed.) (2004) *Reforming Higher Education in the Nordic Countries: Studies of change in Denmark, Finland, Iceland, Norway and Sweden*, Paris: International Institute for Educational Planning, UNESCO.

Ham, C. and Hill, M. (1993) *The Policy Process in the Modern Capitalist State*, London: Harvester Wheatsheaf.

Karran, T. (2005) 'Pan-European grading scales: lessons from national systems and the ECTS', *Higher Education in Europe*, 30, no.1: 7–22.

Leth Andersen, H. (2005) *Eksamensformer: Valg med konsekvenser*, Arbejdspapirer fra Center for Undervisningsudvikling, Aarhus Universitet, No. 2005-01.

Lindberg-Sand, Å. (2005) 'Examination as an enacted view of knowledge: Internal development versus external demands in Swedish higher education', paper presented at AIR annual congress in San Diego, U.S.A., 29 May–1 June.

—— (2007) 'The Swedish Bologna process – reckless race or revitalising reform?', paper presented to the 29th Annual EAIR Forum, Innsbruck, Austria, 26–29 August.

—— (2008) 'The dark matter of the Bologna Process – the phenomenon "Learning Outcomes"', paper presented to the conference Improving Student Learning in Durham, U.K., 1–3 September.

London Communiqué (2007) Towards the European Higher Education Area: responding to challenges in a globalised world, Communiqué from the Ministerial meeting in London, 18 May 2007.

Ministry of Education, Research and Culture in Sweden (2004) *Higher Education in development: The Bologna process from a Swedish view.* (In Swedish only: Högre utbildning i utveckling. Bolognaprocessen i svensk belysning.) Stockholm: Regeringskansliet, Utbildningsdepartementet, Ds 2004:2.

Ministry of Education, Research and Culture in Sweden (2005) *Swedish report: Cooperation in Education and Training under the Lisbon strategy – Towards the 2010 Common Goals.*

Ministry of Education, Research and Culture in Sweden (2006) *National Report for Sweden 2005–2007*, Dnr U2006/9543/UH.

Ministry of Education and Research in Sweden (2008) Swedish report: Ds 2008:13. *A new grading scale* (En ny betygsskala.) Betygsberedningen (U 2007:A).

Norwegian Council of Universities (2000) *Objectives with meaning, (Mål med mening)* En utredning om etablering av en felles nasjonal karakterskala. Agder, Bergen, Oslo, Tromsø, Trondheim – januar 2000. Link to this proposal and links to later papers and works at the link < http://www.uhr.no/utdanning/en_felles_karakterskala > (accessed 31 June 2008).

Selghed, B. (2004) *Ännu icke godkänt. Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen, Avhandling* Malmö: Malmö högskola.

Silén, C. and Harrysson, L. (2004) 'ECTS betygssystem i svensk högskola – ett kvalitetshot?', paper presented to the Conference 'Quality in Swedish Higher Education', Linköpings universitet.

Swedish National Agency for Higher Education (2007) *Granskning av lärosätenas kvalitetsarbete. Anvisningar för självvärdering (Evaluation of Higher Education Institutions. Guidelines for self-evaluation)*, October.

Swedish Government (2005) Government Bill 2004/2005:162 New World – New University. A summary in English of the Swedish government bill.

Wissing, L. (2006) 'Ny karakterskala kræver efteruddannelse til lærerne', *Gymnasieskolen*, 07/2006.