

IMPROVING ASSESSMENT PRACTISE IN NORWAY.

Norway is one of the middle-sized countries in Europe with a population close to 5 millions. School has been mandatory for more than 250 years, and the number of years of schooling has constantly increased. 10 years of school is now compulsory, but we are discussing if an additional three years of upper secondary should be mandatory as well. All students are admitted to upper secondary, no matter how they perform in lower secondary. They are entitled to three to five years of upper secondary, until the end of the school year they turn 21 years of age. Most students graduate from upper secondary after three years, about 50 % from college preparing courses. More than 90 % of these go on to college or university.

During the seven first years of school the children will receive no marks. There is no kind of grading in primary school whatsoever. Grading is introduced in their 8th year, which is the first year of lower secondary. A consequence of this is that daily assessment practises in primary and secondary school are quite different. Let me also add that the concept “assessment for learning” is quite new in Norwegian education. It was first introduced in Norwegian two years ago in a book written by the author of this document and two of his colleagues (Engh et al. 2007).

The Norwegian school system is more centralized than in most other countries. We have had national curricula since 1939, and the regional or local freedom has been very restricted. Text books have been subject to national approval up to recent years, and the subjects of choice have been very few, if any, up til upper secondary.

We also have a short history of national tests. When they were first introduced by the conservative government in 2005, there were so ardent protests among the teachers that they were temporarily withdrawn. In 2007 they were back, but this time they were said to have a formative intension, they were given in the beginning of the school year in stead of the end, and the results were not published openly, so there was no way to produce league tables. As they now were introduced as a tool for the teacher, especially to assist him to teach according to the students’ learning needs, the teachers accepted them. National tests are now given at the beginning of the 5th, the 8th and the 11th school year.

The Ministry of Education, which is situated in Oslo, has an executive branch called the Norwegian Directorate for Education and Training, in this paper referred to as the Directorate. The Directorate has been responsible for analysing students’ outcome on national and international tests and suggesting steps to be taken in order to improve results.

Norway has taken part in the PISA-programme (Programme for International Student Assessment) since the year 2000. When the results of the first tests were published in 2001 the nation was rather shocked to see that our results were no more than average among the OECD (Organisation for Economic Co-operation and Development) countries, and lowest among the Nordic countries. For decades we have believed that our school system has been among the very best in the world, so our politicians' reactions were no less than a shock when the poor results dawned on them. In 2004 and 2007 our students performed even worse, our positions on the international "league" tables were even lower. This development was underscored by results from TIMMS and PEARLS which showed an equivalent situation. Some steps were in need to be taken.

The first step was launched as early as in 2005. To replace the National Curriculum of 1997 the government published a new National Curriculum which stated educational goals for all of the eleven subjects in the National Curriculum. These were described as goals for Competence, and no longer as goals for knowledge, as had been the concept used in previous curricula. The curriculum was presented by a conservative government and was to be enacted in 2006, when the government had already been replaced by a coalition of socialist and liberal parties. Although education has always been an issue of serious dispute between the conservatives and socialists, the new government did not change the National Curriculum and accepted the structure of specified goals for competence for all main topics in the school's subjects. It is however important to understand that all goals are formulated in open terms, using concepts that made it difficult to measure the level of achievement directly. The goals are only specified for four years of primary and lower secondary, that is for the 2nd, 4th, 7th and 10th year. A couple of examples will exemplify how the goals have been stated:

Math – at the end of the 7th year:

The student should know how to develop and use methods for head calculation, estimates and written calculation and use a calculator when reckoning

Language (Norwegian) at the end of the 10th year:

The student should know how to formulate his own opinion in discussions and assess what creates an unbiased argument.

Social studies – at the end of the 7th year:

The student should know how to participate in a discussion of variation in sexual orientation regarding love, relationships and family (Kunnskapsdepartementet 2005).

As the reader will see, these goals are not quantifiable and cannot be measured without extensive discussions that must end in a form of intersubjectivity which can be subject to operationalising. A more prominent characteristic is that these goals can all be achieved to different extents – so that when the teacher assesses any students' work, her feedback can always include the message that the goal has been achieved to some extent. A major responsibility is therefore given to the teacher and her professional assessment as these goals were not created as a foundation for external summative testing.

The history of assessment in Norway is a history of summative assessment and testing. During the school years, individual, written tests with summative purposes constitute a large proportion of the students' school experience. External testing is new, but for the exams, and exams have not been high stakes until the final exams in upper secondary. Up till recent years the class teachers have been responsible for their students' progress and assessment. Even when the students leave lower secondary, only a proportion of them will experience external

exams. Mediocre results on PISA, TIMMS and PEARLS have seriously changed our assessment climate. During a few years in this millennium external testing as well as accountability have become tremendously important among politicians and central school authorities.

The new government was as keen on working to improve the students' learning outcome as the previous one. One of the tasks given to the Directorate was to suggest changes in the assessment procedures in school, and as a consequence the Directorate at the beginning of the school year in 2007 launched a two-year project called Improving Assessment Practise. The project was piloted in 8 primary and secondary schools in August and September of that year, and a slightly revised project started out in December 2007. The project included 77 schools, 59 primary and lower secondary schools, and 18 upper secondary schools. It was scheduled to go on for 18 months, and the schools received financial support from the Directorate for each semester of participation. No schools left the project before it ended in May 2009.

The Purpose of the Project.

The purpose of the project was to trial the implementation of national standards in four subjects, Mother Language (Norwegian), Mathematics, Social studies and Home economics. The overall research question was: ***Will national standards contribute to a more equal and fair assessment practise than the present one?*** Some schools were to trial standards suggested by the Directorate, some were to create their own. The schools were given considerable free space to decide their trialling procedures. There was only one condition for the funding; the schools were obliged to link up with some form of pedagogical expertise. What kind of expertise and the extent of the cooperation or supervision were not specified.

Improving the assessment practise did not mean the instigation of a different paradigm. Studies had shown that the practise of both informal assessment and marking varied a lot between teachers, between schools and between regions, but the variations were not stable from one year to another. Teachers were supposed to assess the students' achievement solely related to educational goals, but studies showed that gender, effort and personal sympathies quite regularly were taken into consideration. The main idea behind the project was therefore to change the assessment practice in accordance with central regulations, towards a more precise and fair summative assessment based on the subject's goals only.

My role in the project was mainly to coordinate the establishment and cooperation of the 77 schools and their owners (municipalities and counties) with the pedagogical expertise. In addition I supervised 11 primary schools in their work with trialling the suggested national standards. These standards were formulated for both high achievers and low achievers. One of the purposes of the project was to develop positive formulations of low achievements, so that the students as well as their teachers would recognise any competence, wether it being high or low. A document to the Parliament issued by the Ministry of Education had stated that the Norwegian assessment culture exerted by the teachers implied telling their students what they did not know and did not master in stead of telling them what they in fact did know and did master. This latter point in combination with applying standards in assessment demanded a change in daily assessment practise which necessarily was to influence teaching procedures as well. The fact that changing assessment practise implied changes in the teachers' teaching methods was an understanding that most teachers gradually developed during the project.

The project was followed by a research team from the University of Oslo. They produced a questionnaire for the project teachers which was to be construed as a pre-test although answered during the third month of the project. Nearly 1000 teachers answered the questionnaire. In addition 49 teachers were interviewed in groups into the fourth month of the project. A corresponding survey (post-test) and focus group interviews took place about four months before the end of the project. Thus the group could report their findings to the Directorate in a final report in May this year.

Each one of the participating schools was to answer a short questionnaire before May 1st. The questionnaire was created by the Directorate and had two main questions. The first was whether the teachers would advise the Ministry of Education to implement national standards, the second was to express their attitude to possible standards, whether standards should function as guidelines or directives.

The word *standard* is also a Norwegian one, but *standard* was not the word used in this project. Instead it was reformulated by using the Norwegian word “kjennetegn”¹ which probably is best translated to English as indicator, but also means something which is typical of the object it refers to. The definition which was assigned to the word “kjennetegn” was said to be a description of quality of the student’s performance, and therefore synonymous with the word *standard*, and slightly different from the concept of *criteria*. In this paper I will use “indicator” for “kjennetegn”, but I will make a point of the word “quality” which is used in the definition of the concept. This points to a qualitative aspect of the students’ learning outcomes which is quite interesting compared to how standards traditionally have been viewed and formulated in other countries (Engh 2009; Glaser and Klaus 1962; Pellegrino et al. 2001).

As the reader will know there have been and still are widespread discussions internationally upon the use of standards in education, since standards were introduced by Robert Stakes almost 50 years ago (ref). The discussion on standards have mostly taken as a presupposition that the standards are formulated in terms that make it possible to measure the outcome either directly or by operationalizing (dividing-splitting) them into measurable magnitudes. In the United States, however, national boards of e.g. mathematics and English have formulated standards in rather open or qualitative statements, but as these standards have been modified by the local state education authorities, the outcomes are very much closed or quantitative, specified in measurable concepts (Engh 2009). A Swedish professor found in his doctoral study that when teachers produced local indicators of achievement, the outcomes were almost unanimously concrete or closed statements, sometimes to the extreme, and as a consequence paid little or no attention to the more open educational goals formulated in the curricula (Tholin 2006).

Many experts on education have warned against the use of standards, Dylan Wiliam, Gordon Stobart, The Assessment Reform Group only to mention a few (see e.g. Stobart 2008). They have been stating that standards tended to make the teachers focus on their students’ performance in stead of their learning. They pointed out that standards limited teaching towards instruction, and that the focus on forming the broad minded student, which should be a central outcome of any educational system, was reduced and replaced by instrumentalistic procedures. Standards have also quite often been linked with accountability, and thereby

¹ Translated directly it means “sign of something recognizable”.

made it more likely to use psychomotoric tests in order to determine the qualities of the school, or, more correct, whether the teaching met the required standards (Langfeldt 2008).

Although these critical perspectives were quite familiar to Norwegian educational researchers, very few of them raised their voices against this project. Highly esteemed researchers put in applications to evaluate the project, a task that eventually was given to our largest pedagogical faculty at the University of Oslo. Only one researcher published her motforestillinger to the project, Professor Astrid B. Eggen (2008), in an article in the journal of our by far largest teachers' union. Her main argument was that the use of standards made it difficult to differentiate teaching and thus reduced the importance of ipsative assessment and teaching in accordance with the individual student's needs.

The standards were formulated by a group of professionals, teachers and employers in the Directorate. The process of formulating indicators that were to show how close to the curriculum's educational goals students had come in their learning is new to Norwegian educationalists. This may explain why the formulation itself created noticeable problems. Some of them were impossible to distinguish from goals, others were too general to function as guidelines for practise. None of them were closed specifications of performance, none were concrete, quantitative or measurable, thus being different from most standards that educationalists from a large number of countries are critical to. Another aspect of the Norwegian way of formulating standards was how several goals were incorporated in one standard. In the National Curriculum one subject can have more than 40 goals specified for one year group, the number of indicators that were meant to include all these goals were usually no more than eight. Some examples of "kjennetegn" may show this more clearly ("kjennetegn" were not suggested for 10th grade):

Math, at the end of the 7th year.

Calculate and estimate with proficiency concerning the choice of methods and approach with and without digital aids.

Language (Norwegian) at the end of the 7th year:

Write fictional and non fictional texts with recognizable genre characteristics and with an appropriate structure.

Social studies - at the end of the 7th year:

Discuss and evaluate how different ways of thinking, values and cultural norms may influence social conditions and human relations both within and outside of Norway
(Utdanningsdirektoratet 2007).

As the project went on, this was confusing to many teachers. Many teachers had been anticipating standards which would assist them in their daily planning; the suggested indicators did not furnish this kind of assistance. The above examples might as well have been construed as goals, they were abstract, left the concretisation to the teachers, and did not specify which goals they referred to. One of the consequences was that several teachers wanted more specified and concrete indicators – a guide to assist them in their daily planning and in selecting student tasks. In other words, they wanted a tool to assist them in their daily, summative practises.

The project started out by giving the participating schools' head masters, deputies and teachers an invitation to a one day congress where they were given information about the background of the project, its purpose and the ideas behind the use of indicators. The leader of

the project, a senior advisor in the Directorate for Teaching and Training, Vivi Bjelke, pointed out, but did not stress, that indicators were not to substitute the goals of the National Curriculum, but should function as guidelines that assisted the teachers to recognize competence as well as the level of competence in their students. (Later it was reported that not all schools and teachers were aware of this perspective, and thus had problems in discriminating between indicators and goals.)

In Norway we have a long tradition of taking the individual student's intellectual abilities into consideration. We have always been proud of the ways our schools are developing inclusion, and we abolished most special schools from as early as in the 1970'ies. The evaluation researchers asked all participating teachers in the first survey whether they thought that the students' abilities and/or their efforts should be regarded when they were given marks, or when they were informally assessed in primary school. Almost two thirds of the primary school teachers and 45% of the lower secondary teachers answered this question positively, although it has been clearly stated in the Ministry of Education's assessment regulations that assessment should be based on nothing but the student's achievement in relation to the subject's goals. In other words, the teachers' view on assessment was not in concordance with the Ministry of Education.

In the Directorate it was believed that it was possible that the use of national indicators would lead to a more equal and fair assessment practise, a practise that took the subject's goals and the goals only into consideration. When the project teachers were asked the above mentioned question again in the second survey one year later, it was of great interest to know whether this attitude had been changed during the project. Were the project teachers now more inclined to focus on the performance and not on the students' abilities and efforts? The result was negative. There had been no significant changes, and the project teachers were as positive to an assessment practise that included the students' abilities and efforts as teachers in a control group.

Generally the use of standards –any kind of standards – is used as an instrument for assessing students' learning outcome, assessment *of* learning. The employees of the Directorate's assessment department were quite interested in anything that could assist teachers in their assessment practise, and "Learning and Teaching in Scotland's" project "Assessment is for learning" had begun to interest more and more employees of the Directorate. Still there was no original intention behind the project to instigate a change in the assessment practise towards assessment *for* learning.

The project proved to have quite an impact on many of the participating teachers as well as the Directorate. As the process went on the teachers' awareness on their assessment practise increased. They gradually became aware of the connections between good teaching practise and assessment. Assessment that in the previous years had been an issue that focused on marking and exams, was raised to be an issue of the utmost importance, not only for the project teachers, but for most people who had a role in the Norwegian education system. In the three decades preceding this project, only one book on assessment had been published. The three last years have seen eight books exclusively on the topic, all of them taking the perspective that assessment is for learning. Assessment is now probably the topic that is most frequently asked for in teachers' in-service training, and primary school teachers, who do not mark their students' work, seem to be as interested in assessment as teachers in secondary education.

Whether focus on assessment has influenced teaching in such a way that it has gained a sustainable change of practise is another matter. A large majority of the project schools answered positively to the main research question. Their practise had shown that standards could function as an aid to a more fair and equal assessment practise, so the majority of the teachers wanted national standards. At the same time they had a distinct attitude to standards as guidelines, not as directives. Generally teachers experience that it is difficult, tedious and time-consuming to teach in accordance with all the goals specified in the National Curriculum. This may be part of the explanation that they are positive to indicators that will guide them in their daily work. Most of them expressed a view on functional standards as more explicit and helpful for their daily planning though. It cannot be concluded that this project has proved that indicators formulated as open and not measurable guidelines will lead to an improved assessment practise which eventually will be welcomed by a large majority of the teachers, nor can it be concluded that national indicators will lead to a greater realization of curricular goals. The side effects seem to me and to a lot of the project teachers, as interesting as the conclusions of the project: A lot of the project teachers are more apt to adopt methods described in literature of assessment for learning. They more often share goals with their students, they discuss indicators of achievement with them, they are more apt to present possible strategies for achieving goals, self assessment as well as peer assessment has become a part of many teachers' assessment practise, and more teachers are apt to more often recognize competence in low-performing students. A typical outcome of the project is that the teachers often have dialogues of the student's achievement with the individual student outside of class, or sometimes in small groups. The Directorate wants this to be a necessary part of the teacher's assessment practise and has included this phenomenon in their suggestions to the Ministry of Education.

The Directorate of Teaching and Training has advised the Ministry of Education to introduce national indicators of achievement in four subjects from this school year on, the subjects being Norwegian, Mathematics, Science and Home Economics. The trend towards a more formative assessment practise was demonstrated in the concluding conference of the project, when Gordon Stobart, a member of the Assessment reform Group, was chosen as the main speaker. The Ministry of Education's regulations on Assessment now include the students' rights to self assessment, and they require all assessment to have formative intentions throughout primary and lower secondary until the end of the 10th, 11th, 12th and 13th school year, when summative grading shall constitute the ultimate assessment. The directives specify the teachers' duty to explain their students how to improve, and the importance of marking has been downgraded. The teachers are also told to have two or more individual dialogues with each student with the main purpose of clarifying the learning goals and discuss how best to get there.

In other words, there is now a national trend towards an *assessment for learning* practise. It was never the intention of the project, it must be viewed more like a side effect, but most likely (and hopefully) as sustainable as the implementation of national indicators. How the Ministry of Education will respond to this is not known when this paper is written, it may be difficult to ride two horses at the same time, one that focuses on National tests and pays serious attention to PISA and TIMSS, and at the same time directs the teachers to change their assessment practise towards *assessment for learning*. The outcome of the Parliament election on September 14th this year will probably be quite decisive.

The story continues; the Directorate suggests that the Ministry of Education shall support, supervise and finance between 100 and 200 schools a year in the next four years to an estimated expense of 20 million Australian dollars in order to improve their assessment practise. But this is an issue for a later conference.

References

- Eggen, Astrid Birgitte. 2008. "Bedre vurderingspraksis" - mot nasjonal standardisering av referanser for vurdering." *Utdanning* 11:58-61.
- Engh, Knut Roar. 2009. "Nasjonale standarder og kjennetegn på måloppnåelse." In *Vurdering, prinsipper og praksis*, eds. Stephen Dobson, Astrid Eggen and Kari Smith. Oslo: Gyldendal akademisk.
- Engh, Roar, Stephen Dobson and Eli Kari Høihilder. 2007. *Vurdering for læring*. Kristiansand: Høyskoleforl.
- Glaser, Robert and D. J. Klaus. 1962. *Proficiency Measurements. Assessing Human Performance*. New York: Holt, Rinehart and Winston.
- Kunnskapsdepartementet. 2005. "Læreplaner."
- Langfeldt, Gjert. 2008. *Ansvar og kvalitet: strategier for styring i skolen*. [Oslo]: Cappelen akademisk.
- Pellegrino, James W., Naomi Chudowsky and Robert Glaser. 2001. *Knowing what students know: the science and design of educational assessment*. Washington, D.C.: National Academy Press.
- Stobart, Gordon. 2008. *Testing times: the uses and abuses of assessment*. London: Routledge.
- Tholin, Jörgen. 2006. "Att kunna klara sig i ökänd natur: en studie av betyg och betygskriterier - historiska betingelser och implementering av ett nytt system." Borås: Högskolan.
- Utdanningsdirektoratet. 2007. *Et Felles løft for bedre vurderingspraksis: en veiledning*. Oslo: Utdanningsdirektoratet.