

**NERA/NFPF 32. Conference****Iceland Pedagogical University, Reykjavik, Iceland**

March 11-13. 2004.

**Olga Dysthe, University of Bergen, Norway**Adr. Seksjon for utdanningsvitenskap, Psykologisk fakultet, UiB, Christies gt. 13. 5019 Bergen,  
Norway**The challenges of assessment in a new learning culture****Abstract**

*The role of assessment has always been crucial in education. It is a truism that assessment shapes how teachers teach and students learn. If we want to know the truth about an educational system, we have to look at the assessment practices and what qualities of students and their work that are being valued by the system. If educational reforms do not include assessment, they often fail.*

Changing perspectives on teaching and learning combined with new demands in society on students' knowledge and abilities are gradually reshaping learning cultures. In such contexts we also need to see assessment through new eyes. While assessment in the past has primarily been a means of certification and accountability, a much wider range of purposes of assessment are now advocated. A new vision for assessment for the new millennium integrates learning and assessment and redefines the roles of students and teachers in the assessment process. Assessment for learning is now gaining importance over assessment of learning. The keynote speaker will discuss new modes of assessment from a socio-cultural perspective, focusing on how to engage students actively in the learning process through for instance portfolios, self-assessment and peer review, but also on the dilemmas arising when formative and summative forms of assessment are mixed.

**Introduction and overview**

The title of this lecture has multiple meanings and I will therefore clarify what I intend to talk about and present some of my conclusions. *A new learning culture*, in my definition, is a culture based on the insights of constructivist, socio-cognitive and socio-cultural perspectives of learning that deals with the changes in society and the new demands on education in the 21<sup>st</sup> century. The basic premise for this lecture is that a new learning culture is emerging and that it needs alternative assessment forms in order to develop. I have chosen to organize my lecture around five challenges, with different addressees:

The first challenge is the 'wash-back effect' of assessment, meaning that assessment has a strong influence on how teachers teach and how students learn. This alone makes assessment a crucial topic. Assessment practices are difficult to change, but if they remain unchanged, important aspects of a new or emergent learning culture are in danger.

The second challenge comes from developments in society and the ensuing expectations of what knowledge and competence students have acquired while studying. My claim is that alternative assessment forms are needed to meet these expectations.

The third challenge is the importance of alignment or congruence of theories of learning and

assessment. Today we experience a gap between theories underpinning a new learning culture and the tacit theories underlying traditional assessment practices.

The fourth challenge is how teachers and students, who are socialized into traditional assessment practices, deal with the new modes of assessment. I will use an empirical study of portfolio assessment in Norwegian teacher education to illustrate some dilemmas as experienced at grass root level.

The fifth and last challenge I want to focus on is how to meet the strong international trend of measurement and accountability, and I will argue that unless alternative assessment practices are theoretically grounded and quality secured, and unless educational researchers and teachers join forces, the strength of the international testing regime may sweep alternative assessment away and bring the new learning culture with it.

Because I have been working in the interface between writing research and education for many years, I have chosen an example from writing research to frame this lecture. I have noticed that even though writing research is concerned with issues of development, learning and assessment, education research and writing research seem to function as quite separate research communities with their own conferences and journals, and the knowledge created in one area is surprisingly unknown to researchers in the other. I will also exemplify my points by referring to recent studies of portfolio assessment in Norway.

### **I. The challenge of the wash-back effect of assessment: a historical example**

The strong influence of assessment on teaching and learning is much talked about but rarely documented. I would like to start by telling you the first part of a true story about the transformation of a learning and assessment culture. It is based on first hand experience, but it has also been published in the last issue of the journal *Assessment of Writing*.

In 1985-86, on leave from my job as a teacher in the Norwegian secondary school, I spent a year in southern California, where I visited classrooms all over the San Diego School district. I had a small grant to study the use of computers in schools, but I ended up studying the way they taught and assessed writing. I had discovered the writing process movement and happened, by sheer luck, to be caught up in the midst of an incredible period of educational change. In the mid 1980s the state of California challenged the pervasive testing culture in the United States, and became a progressive innovator in assessment development and in the professional development of teachers. The California Assessment Program (CAP) had initiated "the most ambitious achievement test in writing in the nation, a direct assessment of writing that evaluated a variety of genres and forms" (Murphy 2003, p 24).

The CAP was the product of a broader movement toward performance assessment that had begun earlier. Politicians had been convinced by educational researchers and teachers that the previously dominating multiple-choice tests did not tap into the full range of students' knowledge and abilities and were to be replaced by alternative assessment formats.

California Writing Project, well known throughout the whole of the US, and also in Scandinavia, supported the change and provided staff development, directly linked to the writing assessment. Scoring guides were part of the widespread support material that was produced. The reformers worked from a pragmatic position, formulated by an American colleague: "If you can't remove the testing system, and if teachers will always teach to the test, our job is to create tests worth teaching to."

I happened to come in the middle of all this and experienced what in effect was a transformation of a learning and assessment culture in the largest of the American states, with a

population of more than 20 million. When I came back in 1990 as a doctoral student at the University of California, Berkley, writing research was burgeoning, portfolio assessment was being introduced in writing classrooms, there was a shift in the predominantly cognitive theory base, with its focus on the individual, towards a socio-cognitive and socio-cultural foundation for writing theory and pedagogy. I will return to the second part of this story, which does not have a happy ending, but let us consider *what lessons can be drawn from the California story*:

- First of all it is encouraging because it tells us that learning and assessment cultures *can* change.
- Second, it confirms that assessment is “the tail that wags the dog”. There is no doubt that the change in assessment became the engine of the change process, and a strong engine was needed to move this large state, where the measurement culture was just as entrenched as in the rest of the US.
- Third, it documented that educational researchers deeply committed to classroom practices in coalition with teachers represent a strong force. The change was teacher driven, but needed the theoretical anchoring in writing and learning theory provided by Harvard scholar Janet Emig, Linda Flowers at Carnegie Mellon, Shirley Brice Heath at Stanford, Charles Cooper at UCSD, Ann Dyson at Berkeley and many others who provided landmark research studies underpinning the new practices.
- Fourth, new assessment forms involve a major conceptual shift and both teachers' and students' need support to change their practices. We must be alert to the danger of trivialized and subjective assessment practices under the guise of new assessment, if the changes are not anchored.

The driving force behind the innovation of the teaching and assessment culture in California was the conviction that students needed a much broader writing competence to meet the demands of their future lives. This meant a change from just teaching and testing writing as a decontextualized skill to focusing more authentic forms of writing. Schools and universities have a duty to prepare their students for the demands of the future, and this is also a challenge to test makers. I will therefore in the next section briefly discuss some challenges of the 21<sup>st</sup> century.

## II. The challenge from developments in society

### 1. What knowledge, skills and experiences do students in the 21. century need?

When the role of student assessment is changing today, it is “largely because today's students face a world that will demand new knowledge and abilities, and the need to become life-long learners in a world that will demand competences and skills not yet defined” (Segers, Dochy, Cascallar 2003, p.1). People at all times have been ‘lifelong learner’, but today the term takes on a new meaning because of the rapid changes. Our students need to learn how to learn because the knowledge and the skills they have acquired in schools and universities will be outdated or not sufficient. Students will still need solid basic knowledge, but they need also to develop a broad range of competences besides disciplinary content knowledge. The society of tomorrow will require people who are flexible and able to continue to acquire new knowledge and learn new skills. All this is common knowledge now.

The information age is characterized by a steadily growing, dynamic and changing mass of information. Students need digital literacy, but also a variety of competences in order to function well in the information society. Birenbaum (1996) has analyzed and categorized these competences and skills in the following way:

a) cognitive competences such as problem solving, critical thinking, formulating questions, searching

for relevant information, making informed judgements, efficient use of information, conducting observations, investigations, inventing and creating new things, analyzing data, presenting data communicatively, oral and written expression; b) meta-cognitive competences such as self-reflection, or self-evaluation; c) social competences such as leading discussions, persuading, cooperating, working in groups, etc. and d) affective dispositions such as for instance perseverance, internal motivation, self-efficacy, independence, flexibility, or coping with frustrating situations (Birenbaum, 1996, p. 4).

It can be discussed to what extent these competences are actually new or more important than before. The main point, however, is not the newness, but the fact that a learning culture for the 21<sup>st</sup> century must help students develop these competences as an integral part of acquiring content knowledge. In higher education there is an increasing awareness that traditional lecture based teaching and exams are insufficient to meet such goals, and at the same time the previous strong division between learning in school and learning at work is becoming less clear.

In Europe OECD has conducted a large scale investigation among university graduates, employers and academics in 16 countries to establish a list of the most needed competences in higher education (also called 'capabilities', 'generic attributes', 'key skills' or 'key learning outcomes'). The result is as follows for all the three groups:

1. Capacity for analysis and synthesis
1. Capacity to learn
1. Problem solving
2. Capacity for applying knowledge in practice
3. Capacity to adapt to new situations
3. Concern for quality
4. Information management skills
4. Ability to work autonomously
5. Teamwork
6. Capacity for organisation and planning

Interestingly, 'teamwork' is much higher on the list of employers and university graduates, than among academics!

It should be obvious to most people that 'back to basics' and traditional testing is not the way to foster these competences, - a new learning culture is needed. Introducing new technology is not in itself a solution, because unfortunately educational programs that utilize information technology, too often digitalize oldfashioned knowledge instruction (Engelsen 2003).

And if "we get what we assess", how do we assess competencies? Should we follow Australia, for instance where they have developed a compulsory standardized national test for all university students, called "the Graduate Skills Assessment Test"? Australian education specialists warn against this, and claim that it will result in instrumentalism and reductionism to impose a decontextualized tests in areas where sensitivity to context and culture is important (Clerehan et al,

2003).

In the US education authorities plan to move in the same direction. This has prompted the highly respected Stanford professor, Richard Shavelson, who by the way is a specialist in psychometrics and testing, to protest against the narrowing of competences to those that can be tested. In an article in the journal *Change* he discussed assessment in relation to the broad range of cognitive, personal, social and civic goals which the higher education institutions themselves promoted. The cognitive outcomes most often cited by institutions were:

A. learning skills ... and knowledge of particular facts in a discipline ....; B) reasoning with or applying knowledge to solve problems .... C) learning about one's own learning so as to monitor progress in problem-solving and to learning in new situations (Shavelson & Huang, 2003 p. 12).

The personal and social outcomes were also reflected in the public's perceptions of the purposes of higher education. A national poll ranked higher education's goal in the following order (Immerwahl 2000):

- Sense of maturity and [ability to] manage on [one's] own (71%)
- Ability to get along with people different from self (68%)
- Problem solving and thinking abilities (63 %)
- High-technology skills (61%)
- Specific expertise and knowledge in chosen career (60%)
- Top-notch writing and speaking ability (57%)
- Responsibility of citizenship (44%)

Shavelson and his colleague make the point that in debates about learning, assessment and accountability the goals of higher education are almost always confined to cognitive output measures and that those who advocate large scale tests, don't realize the narrow scope of what is actually possible to test. His advice is to assess a broader range of what we value, not just the cognitive outcomes of education.

To repeat my major point: new modes of assessment are needed both in order to foster and to assess this broad range of desirable knowledge and competences. And now it is time to take a closer look at what characterizes new modes of assessment.

## 2. What are characteristic traits of new modes of assessment?

'Alternative assessment', 'direct assessment', 'authentic assessment' and 'performance assessment' are all terms used more or less synonymously. The term "alternative" is relativistic, as it depends on what the phenomenon in question is an alternative to. Alternative assessments are alternatives to the testing and measurement tradition; a shift from so called objective, short answer tests, for instance fill-in-blanks, multiple-choice, true/false. These are being replaced by various forms of performance assessments, project based assessment, profiles of student learning, portfolio assessment, just to mention some. We see changes in purpose, substance and forms of assessment. Gielen, Dochy & Dierick (2003, p 37) have identified key factors:

- New modes of assessments are *contextualized* ('authentic') instead of

decontextualized.

- They aim at assessing much *broader aspects of learning*, and describe the student's competence by a profile often based on *multiple sources and over time*, instead of by a single mark.
- The purpose of assessment is primarily learning, not selection or certification, hence the slogan "*assessment for learning*" instead of "assessment of learning".
- There is an *integration of assessment in the learning process* instead of separating the two. Assessment as a tool for ongoing learning is increasingly stressed.
- There is a shift in assessment *from low levels of competence to higher levels*. According to CRESST (National Centre for Research on Standards and Student Testing at the University of California, LA), the core types of learning that should be assessed, are conceptual understanding, knowledge representation, problem solving, communication and team work. This means a shift in focus from assessing the reproduction of knowledge to higher order skills.
- *Social, affective as well as metacognitive aspects of learning* are included in assessment. The rationale for this can be found in the shift from a unidimensional to a multidimensional view of intelligence (Gardner 1984, Seeger, Dochy & Cascellar 2003). Intelligence is more than cognition. "Only if we expand and reformulate our view of what counts as human intellect will we be able to devise more appropriate ways of assessing it and more effective ways of educating it." (Gardner, cited from Davies & LeMahieu 2003, p.141). Others would argue from a socio-cultural perspective that because learning is fundamentally a social activity, it is vital to assess how students participate, contribute and take advantage of interactions.
- The responsibility for assessment is shifting to the student. *Peer review and self assessment* are often included, the latter both as a means and a goal.

Taken together these changes amount to a paradigm shift in assessment, which goes far beyond the changes I have just described in writing assessment in California, where the changes were within the confines of a large scale testing system.

Before I continue, I would like to make clear that I am *not* against all kinds of testing. I do believe we need *multiple* forms of assessment, both because of the complexity of educational goals and because we cannot avoid *selection*. But my topic today is the new modes of assessment.

### III. The challenge of aligning assessment with theories of knowledge and learning

#### 1) Changing paradigms

Three well known British and American researchers, Patricia Broadfoot (1994), Caroline Gipps (1999) and Lorrie Shepard (2001) have examined the epistemological underpinnings of both traditional and new forms of assessment. In separate studies they have come to the conclusion that the testing culture is aligned to outdated theories of learning, and that new modes of assessment are needed today to fit and support new ways of teaching and learning. Shepard underlines the importance of knowing where traditional views of testing came from and how closely they are connected with past models of curriculum and instruction (p. 1067). In a very interesting article called "The role of classroom assessment in teaching and learning" Shepard has shown how the dominant paradigm in the last century was a combination of behaviorist learning theories, hereditary theories of intelligence and curriculum theories and scientific measurement.

Shepard argues that this alignment has been broken because constructivist and sociocultural perspectives on learning created a new theoretical foundation for teaching, while the *thinking* about assessment and the traditional test forms were carried over from the previous paradigm. What she tries to illustrate in the figure below, is that from the 1980-ies there has been a conflict between *old* assessment forms and *new* views of learning. During the last ten years, however, she claims that a new paradigm is emerging where assessment and learning theories gradually are aligned once more in what she calls “classroom assessment”.

Shepard discusses how dominant theories of the past continue to affect current practices and perspectives and she underlines the importance of understanding the origin of as well as the strength of the measurement thought patterns, and the extent to which they still dominate over ways of thinking:

Belief systems of teachers, parents and policymakers are not exact reproductions of formal theories. They are developed through personal experience and forms popular cultural beliefs. Nonetheless formal theories often influence implicit theories that are held and acted upon by these various groups. Because it is difficult to articulate or confront formal theories once they have become a part of popular culture, their influence may be potent but invisible long after they are abandoned by theorists (Shephard 2000, p 1068).

To Shepard, learning theory is primary and assessment practices should follow. For instance, when learning according to constructivist theories is an active process of sense making and not a passive reception of knowledge, assessment tasks must ask for production, not just reproduction of knowledge. One of the reasons why portfolios have become so popular is that they provide the space needed for such assignments.

## **2. New modes of assessment aligned to cognitive and situated views of what it means to know**

In the landmark article “Cognition and learning” in 1997, James Greeno, Allan Collins and Lauren Resnick outlined some basic differences between the three main views of knowing and learning: behaviourist, cognitive and situated or sociocultural, and they have formulated succinctly some important differences in views of assessment:

The traditional behaviorist perspective supports a quantitative view of knowing and learning, in which assessment involves independent samples of knowledge or skills to estimate how much of the domain a student has acquired. The cognitive view of assessment emphasizes questions about whether students understand general principles in a domain and whether they use methods and strategies that are useful in solving problems in the domain. The situative view of assessment emphasizes questions about the quality of students' participation in activities of inquiry and sense-making, and considers assessment practise as integral components of the general systems of activity in which they occur (Greeno, Collins Resnick 1997, p. 37).

Tests of ability or knowledge composed of atomistic items make sense from a *behaviourist perspective* when the question we need to answer is some version of “How much?” i.e. How much does a student know in some domain such as mathematics, history, or biology? But when from a *cognitive perspective* knowing is seen as understanding general principles in a certain domain and the ability to reason and to solve problems, alternative assessment forms are needed where students through projects or other extended performance can demonstrate their ability to understand, reason, solve problems and communicate.

“*Performance assessment* provides a bridge between the cognitive and the situative perspectives on learning, because the extended performances needed to assess conceptual understanding and reasoning often also involve engagement with other people and with tools and artefacts that create natural, or “authentic”, situations of activity” (Ibid, p. 39). This is one reason

why there are many points of agreement between proponents of the cognitive and the situative perspective on learning when it comes to assessment. Another reason is their common need to fight the still so pervasive measurement paradigm of assessment.

In the situative or sociocultural perspective knowing in a domain is closely associated with the ability to participate in communities of practice, for instance disciplinary communities like biology, history or literature or professional communities like teaching, engineering or nursing, with their particular activities and discourses. It follows logically that assessment should be contextual and closely bound up with the normal activities in the community of practice, whether this is a classroom or a workplace, hence "classroom assessment" and 'authentic tasks' (Lave & Wenger 1991, Wenger 1998, Wenger, McDermott & Snyder 2002). And since a characteristic of a community is that its members are dependent on one another, collaborative and collective tasks are combined with individual tasks.

It also follows from a situative or sociocultural perspective that students are not just *objects* of assessment, they are *participants* in all aspects of assessment, from the formulation of criteria and quality standards to creating meaningful assignments and assessing their own and other students' work. This is an integral part of the socialization process, of learning what it means to know in a particular domain and what it involves of qualifications and responsibilities to be for instance a biologist, a teacher or a nurse. This is formative assessment in its true sense, and it includes the formation of identities.

Theories of learning do not only deal with epistemology, but include ontology, i.e. what it means for somebody to be. "Learning involves not only becoming a member of a community, not only constructing knowledge at various levels of expertise as a participant, .... Learning entails both personal and social transformation" (Packer & Goicoechea 2000, p. 228). There is greater awareness today of how historical and cultural factors constitute or form **identities**, images of possible selves, and how important this is in learning. Assessment plays a crucial role in identity formation. When students are assessed, whether by a teacher or by peers, the very language that is used is important for how they gradually view themselves as learners. The individual evaluates himself or herself in the light of how others judge them. "Students perceptions of their abilities are particularly sensitive to social comparison of information" (Gipps 1999 p. 383). When fostering identities is seen as an important aspect of learning, it is necessary to design assessment forms that give space for students to express themselves and their emerging identities. One of the reasons why portfolios have become so popular in basic schools is probably that it offers pupils such a space.

Here is also a connection to *theories of intelligence*. While hereditary theories focused primarily on students' inborn abilities, the focus changed to how cognitive abilities could be developed through socially mediated learning opportunities (Shepard 1974). Earlier narrow views of intelligence have been replaced by notions of *multiple intelligences* (Gardner 1984). When the need to foster and value a wide range of creativity as well as social and emotional development is foregrounded, not just scholastic aptitude, this demands new modes of assessment. One of the earliest and most successful portfolio projects in the US, Arts Propel, was directly inspired by Howard Gardner's work, and Lars Lindström's work on alternative assessment in Sweden, has followed the same line, as have several Danish schools.

#### **IV. Challenges to teachers and student in practicing assessment for learning: Examples from the Nordic countries**

##### **1. Portfolio assessment - an introduction**

Because of the growing interest in portfolio assessment in the Nordic countries, I will use portfolios as an example of what challenges new forms of assessment may involve for students and teachers. Portfolios may be used for different purposes, and it is the purpose that determines what is collected,



who collects it, how it is collected, who looks at it, how they look at it and what they do with it (Dysthe 1996, Klenowski 2002). Both purpose and use show great variations across school sites. Two of the most quoted definitions of portfolios focus on the portfolio as a purposeful collection of evidence of learning and progress over time. The first of these definitions focuses on *collection, reflection and selection*, and emphasises students' active role, while the second is less specific:

A portfolio is a purposeful collection of student work that exhibits the student's efforts, progress, or achievements in one or more areas. The collection must include student participation in selecting contents, the criteria for judging merit, and evidence of the student's self reflection (Paulson, Paulson & Meyer (1991p. 60).

Portfolio assessment is a purposeful, multidimensional process of collecting evidence that illustrates a student's accomplishments, efforts, and progress (utilising a variety of authentic evidence) over time" (Gillespie et al 1996 p. 487).

There are many types of portfolios, for instance: "Showcase portfolio" (a pupil's choice of favourite products), "Development portfolio: a collection chosen to show development during in a domain over a certain period. "Learning portfolio" focuses on the learning process, while "Assessment portfolio" contains documentation of a particular course for summative assessment purposes. Lindstrøm (2004) talks in addition about "Project or process portfolio" which contains finished product(s), for instance in art or crafts, as well as a logbook showing in detail the process of producing it. "Employment portfolio" resembles an extended CV and is used for job application. As I see it, the biggest challenge to schools that want to use portfolios, is to determine what its purpose should be.

In higher education portfolios and other forms of documentation involving reflection and self assessment have their origin in one of two very different traditions. One is the competency movement which is basically instrumental in its approach, and the other is humanism and constructivism. The focus of the first one is on documenting competence in relation to detailed learning goals. One notable example of this is the European Language Portfolio based on the Common European Framework of Reference, with its detailed rubrics for self assessment of competency in a foreign language. The danger of instrumentalism exists unless the external standards and norms are internalized and integrated with internally developed goals (Brown, Bull & Pendlebury 1997, p 185). When they are appropriated and contextualized in a reflective manner, they can, however, have an innovative effect on pedagogy, as can be seen in several Nordic countries.

When portfolios have their origin in humanism and constructivism, the aim is to enhance disciplinary and personal development. Reflection often plays an important part together with peer review and teacher response. The purpose is to develop new levels of understanding. Teachers in primary and in professional education are particularly attracted to this approach. According to Brown, Bull and Pendlebury (p.186) these two approaches to portfolios have created confusion both in the literature and in practice.

## 2. Some Nordic examples of portfolio studies

Portfolios have been used by individual teachers over the last ten years from preschool to university, but research studies are still relatively few. The first portfolio project to be initiated by educational authorities in Norway was a writing portfolio project at Grønåsen ungdomsskole from 1993-1996 (Dysthe et al 1996). It contained central characteristics of portfolio assessment: the pupils were actively involved in the assessment process through a) systematic use of peer response, b) criteria discussion and c) self evaluation. Besides, *celebrating their strengths* and *formulating specific areas of improvements* were new and challenging elements for the pupils who participated in the development of a new learning culture through portfolio work. Both these aspects have an empowering effect on students at all levels of education.

In Finland a group of researcher headed by Linnakylä and Kankaanranta at the University of Jyväskylä has explored portfolios in preschool and elementary school, especially digital portfolios (Kankaanranta & Linnakylä 1999). Several research projects, ranging from kindergarten to university, are in different ways connected to portfolios as a tool for learning, assessment and organisational development. Their research has focused on portfolios and personal study plans, assessment of portfolio thinking, content analysis of portfolios and feedback, quality and assessment criteria, different uses of a portfolio: purpose, reflection, collaboration and interaction. Marja Kankaanranta has researched how teachers communicate and collaborate through networking and digital portfolios and shown how portfolios build bridges between the children's successive learning environments from kindergarten to school and thus contribute to continuity in the childhood learning and teaching. Finland has also involved several universities in a three year trial of the European Language Portfolio.

In Denmark Mi'janne Juul Jensen has tried out portfolio assessment of writing in a Copenhagen gymnasium over several years, followed up by Ellen Krogh. The project, which was inspired by the Grønåsen project, is documented for the Danish Ministry of Education (Juul Jensen 2000, 2003, Krogh & Juul Jensen 2003). The portfolio does not replace the final exam, but forms the basis not for students' test grade, but their school grade. There are also Danish portfolio projects in teacher education, and Ålborg University Elsebeth Korsgaard Sorensen at has studied virtual portfolios (Takle, Sorensen, Taber & Fils 2000).

In Sweden, the concept of portfolio assessment appears in the literature from the early 1990-ies (Lindström 1992) onwards, inspired by new modes of assessment in arts education (Linds-tröm 1994), creative writing (Molloy 1996) and general education (Ellmin 2003). Although portfolios today are used in practically every school in the country, very few research projects have studied the potential benefits of using this instructional tool.

In a National Agency for Education study portfolios were used to follow the progression of young people's creativity in the visual arts from preschool to upper secondary school (Lindström 1999). The study raises important questions: Is school over-emphasising products on behalf of processes, knowledge on behalf of learning? Will portfolios, by attending to students' performance, help to promote more thoughtful ways of learning and problem solving? In a later work Lindström (2004) used portfolio assessment in conjunction with Kelly's repertory grid technique to investigate conceptions of craftsmanship among teacher educators and artisans. Large similarities were found in the way teacher educators and artisans assessed craft portfolios, the craftspeople being product-oriented while the teachers were more process-oriented.

In Norway there have been a number of portfolio projects in schools since 1993. In higher education the interest in portfolio assessment exploded after the introduction of the Quality Reform, where the integration of teaching and assessment was advocated and portfolios mentioned as an example of alternative assessment forms (Parliamentary Proposition 27/2001). Most of the studies focus on portfolio assessment in teacher education (Allern 2003, Dysthe 2003, Dysthe & Engelsen 2002, 2003, 2004, Hauge 2001, Hauge & Wittek 2002, Ludvigsen & Flo 2002, Otnes 2002, 2003, Wittek 2002, 2003) and nurse education (Havnes, Taasen, Jensen & Lauvaas 2003, Wittek forthcoming). Raaheim (2000, Raaheim & Raaheim 2003) Dysthe & Tolo (2003) has investigated the use of digital portfolios at the University of Bergen, while Wittek (2003) reports from the University of Oslo.

## 2. The project "Alternative Assessment in Teacher Education"

The project was funded by the Ministry of Education as part of a larger project on the use of Information and Communication Technology and the portfolios were digital. Three institutions were involved, the Department of Teacher Education at the University of Oslo, the University College of Vestfold and the University College of Stord/Haugesund. Portfolio assessment was used in

mathematics, Norwegian language and literature, natural science, religion, pedagogy and subject didactics, and a great variety of models were tried out. A general finding was that a great majority of students were very satisfied with portfolio assessment.

Many of the challenges of portfolio assessment as we saw them were related to what Shepard called “the emergent paradigm”. Both teachers and students have educational experiences rooted in an ‘old’ learning culture, and the change from traditional assessment to portfolio assessment challenged many of the taken-for-granted ideas and practices. One of these ideas is that grades are supposed to show what students can do alone, and not as Vygotskij advocated, how well they could utilize the help of others. Traditionally teachers make assignments and students solve them; teachers decide criteria and assess products, just to mention a few practices and roles that are no longer obvious when assessment is integrated in the learning process. The three first challenges I have chosen to highlight reflect very fundamental issues, the two last ones more specific issues related to the competences both teachers and students need in order to utilize the potentials of digital portfolios.

### *1. Balancing formative and summative assessment*

Just to illustrate one dilemma of formative and summative assessment: Students got extensive peer and teacher feedback on their portfolio assignments and they appreciated this very much. They were not in doubt about the great learning potential. But there was also a danger that that summative assessment might invade all aspects of their work, as well as the more specific danger that extensive teacher guidance and peer collaboration would raise the question when it came to grading the portfolio: “Whose portfolio is it really?”

Our sites also experimented with different summative assessment formats, for instance: a) grading the portfolio alone, b) oral exam based on the digital portfolio, where the student would present the portfolio and discuss it with the exam commission, c) written exam based in various ways on the portfolio work.

### *2. Balancing teacher and student control - new roles*

Teacher education in Norway has traditionally had a strong focus on teacher and curriculum control. Portfolio assessment provides space for students to take a greater responsibility in creating and demonstrating their knowledge. When the old traditions prevailed, portfolios might end up just as a collection of very traditional assignments, when responsibility was shared between teachers and students in productive ways, learning abounded. Giving over control to students was often a difficult process, as it involves new roles for both parties. It takes time to find the right balance.

### *3. Balancing individual and collaborate work*

Traditional assessment focuses on individual performance. Even though all three sites encouraged extensive collaborative learning processes around portfolio assignments, with a few exceptions, the portfolios were individual, with some group projects included. A major challenge is how to handle the dilemma of collaborative work and individual assessment.

### *4. Digital portfolios - “filing cabinet or learning arena”?*

Hildegunn Otnes, one of our project team, raised the question of whether the digital portfolio functioned primarily as “a filing cabinet or as a learning arena”. Digitalized portfolios were obviously a good administrative tool, but in many cases the specific learning potentials in the digital medium were not utilized. Some of the issues at stake were:

- Transparency, which varied on a scale from the portfolios only being accessible by the teacher and the student to being open to the public at large on WWW

- Hypertexts and multimedia `texts' or only linear, written texts?
- Collaboration across sites?

*Example 1: The Norwegian Novel Project*

A joint portfolio assignment in Norwegian literature at Vestfold and Stord University Colleges asked students to analyze a number of selected modern novels, produce collective analytic hypertexts and publish them on the web. Students from the two sites were placed in cross-institutional groups. They co-constructed shared web-sites and hypertexts about the novels. They used chat for brain-storming, peer response and to discuss where to make the hyperlinks. E-mail was used for more extensive text responses. This was an example of students creating new knowledge by utilizing the distributed knowledge of their fellow students.

*Example 2: Nordbygda School Project (Nature, environment and Pedagogy)*

## **5. The challenge of “giving the students the competences they need to succeed”**

New competences:

- Giving & utilizing feedback
- Formulating, negotiating and using quality criteria
- Reflection
- Peer assessment
- Self assessment
- Digital literacy and competence in netbased collaboration
- How to deal with sources (knowledge & ethics)

*Summary of challenges*

*For the institution*

Our teacher education project illustrated clearly that portfolio assessment *has* the potential of changing the learning culture and to act as bridges between theory and practice, but it does not follow as a matter of course. It depends on the extent of participation by the students in all phases of development of a portfolio culture. It also depends on whether there is institutional support for integrating portfolios in the community of practice, and not just tolerance for individual teachers' innovative practice. Some practical challenges for the institution are:

- to establish clear frameworks and guidelines for portfolio assessment
- to be clear about organizational details and procedures, including procedures when students fail
- grading practices and how to deal with reliability and validity issues
- how to develop an interpretative community for assessing portfolios

*For teachers:*

- how to balance control and freedom in the portfolio process
- how to integrate process and product in the assessment
- how to deal with the dilemma of formative versus summative assessment
- how to assign and assess reflective texts
- how to involve students in criteria discussions
- how to promote self assessment
- how to utilize the particular potential in the digital medium
- how to develop new teacher and student roles

*For students*

- written and oral communication skills
- ICT competence, collaboration skills and peer review skills
- time management
- the quality of feedback
- explicit criteria
- metacognitive strategies
- new roles in the learning and assessment process
- freeriders, plagiarism

This study tells us that the introduction of portfolio assessment highlighted the need to develop many of the competences or skills that Birenbaum (1999) identified as crucial for the 21<sup>st</sup> century. But we draw the conclusion of our empirical material that such competences should not be taught or assessed in isolation, but have to be integrated in disciplinary content work. This is also in accordance with a sociocultural perspective on the deeply contextualized nature of all activities. In order to exemplify what this entails I will discuss in more detail one of the many challenges I have listed, namely the issue of peer response and explicit assessment criteria.

#### **4. Peer response and explicit criteria -important elements of portfolio work**

In the wake of new modes of assessment students have gained new roles in the assessment process for which few of them have been trained. Students are being engaged in peer response, peer assessment and self assessment, but more often than not without being prepared for these new tasks. While 'peer response' or 'peer feedback' is a term used particularly in process writing and is per definition formative and often group based, 'peer assessment' can be formative or summative. It can be argued that assessment is always an element in giving feedback, however informal it is, as there will always be tacit or explicit norms underlying the comments given. It is therefore not surprising

that students demand explicit criteria which they believe will help them give peer feedback as well as help them understand the background for teacher assessment of their own work. This is a reasonable assumption. In the assessment literature there has been a consensus that explicit criteria is a basic requirement for a well functioning assessment system, and sometimes you get the impression that this is a panacea that will solve the quality problems of new forms of assessment (Biggs 1996, Wittek & Dysthe 2003).

The issue is, however, rather complex. On the one hand it is too simple to think that “if all were made explicit this would be sufficient to establish standards for all subjects” (Rust et al 2003, p. 148), and on the other hand it is unsatisfactory to rely totally on tacit criteria: “I know a good paper when I see one”. Research conducted at Oxford Brooks University indicate that even carefully worked out criteria were “of limited practical use if presented in isolation without the benefit of explanation, exemplars and the opportunity for discussion” (Ibid p. 151). The researchers concluded that criteria understanding and use is dependent on a combination of tacit and explicit knowledge:

... without active involvement through discussion and debate, the development of a common view on standards and level is problematic, if not impossible - even within a close-knit community of a single academic department. ... Polanyi (1998) and Tsoukas (1996) among others, argue that the transfer of useful knowledge involves the transmission of both explicit and tacit knowledge. Consequently, a single-minded concentration on explicit knowledge and careful articulation of assessment criteria and standards is not, in itself, sufficient to share useful knowledge of the assessment process (Ibid).

From a situated or sociocultural perspective the embedded nature of quality criteria is obvious and the challenge is finding good ways of socializing students into the particular disciplinary or professional community's understanding of what counts as quality performance, whether it is good practice or good text in a certain genre. In the Grønåsen writing portfolio project this was done systematically for instance through regular discussion of exemplars of texts in different genres, both written by students and by professionals (Dysthe 2001). At Oxford Brooks University they designed workshops for Business students in order to discuss criteria and train grading student papers. In a controlled experiment with a large number of participants they found that those students who had participated in a 90-minute workshop where they had to grade student papers and discuss criteria both among themselves and with teachers, not only improved their ability to peer assess, but also their own performance. The control group performed considerably lower, even though those students also had access to the criteria set but had not been involved in the discussion (Rust et al 2003).

Another study from the Netherlands shows similar effects of training teacher education students in peer assessment (Sluijsmans, Brand-Gruwel & Merriboer 2002). The findings of these studies are not surprising to teachers but they may have considerable implications in pedagogical practice in higher education, particularly at universities where many disciplines have no tradition for integrating training of competences and skills. In a study we made of digital portfolios in a traditional university course, students were required to give written, netbased feedback to fellow students' argumentative papers, and 6 feedback texts pr student were submitted in their portfolios and included in the basis for summative assessment (Dysthe & Tolo 2003). Our findings, however, indicated that criteria were not discussed, that the quality of the peer feedback was very variable and that students did not use the feedback for revision purposes. This is just an indication that there is considerable improvement potential in the area of peer response or peer assessment and formative use of criteria.

The problems of peer response and the use of criteria which came to the fore in our empirical study of portfolios in teacher education in Norway, echo in many ways the concern and issues I became involved in during my California stay that I told about in the introductory story. It is time for the not-so-happy-ending.

## **V. The challenge from global trends towards testing and accountability systems**

## 1. The story of writing assessment in California revisited

The progressive writing assessment program from the mid 80ies to the mid 90ies has been reversed. A radical change has taken place in assessment policy in California characterized by high-stakes accountability and measurement. High stakes testing was introduced to monitor and measure student learning, to control and compare schools and to drive curriculum development and selection (Murphy 2003, p 27). California's new STAR program, (Standardized Testing and Reporting) introduced in 1998 a multiple-choice test of students from Grades 2-11 in reading, writing and mathematics. The primary means for assessing writing in high school was once more multiple-choice.

The effect of this on the teaching of writing has been researched by Sandra Murphy. The general tendency is that while teachers in 1988 were concerned with students' overall ability to write good texts in a variety of genres, in 2001 teachers emphasized the teaching of isolated skills and grammar. While 88 % of the teachers in 1988 said that they put a lot of emphasis on genre characteristics when responding to student writing, only 56 % did so in 2001. It actually seems that the new test has managed to destroy much of the progress that was made in writing pedagogy over the last 20 years. Murphy's investigation also shows the importance of creating tests worth teaching to. In 1987 as many as 94 % of the teachers in the survey answered that the new writing test program would improve their teaching, while only 26 % of teachers in the 2001 survey said the same thing about the new STAR test.

It is a rare situation that assignment paradigms change so dramatically in such short time and Sandra Murphy's study of the impact on curriculum, teachers and teaching has got considerable attention.

## 2. Looking ahead in the Nordic countries

The demand for high quality educational services world wide has created a generalized picture of educational systems in crisis, and the international trend is to deal with this by introducing national assessment system, with all the trade marks of the measurement and testing culture I have talked about before. These systems are characterized, as we all know, by the use of performance indicators, concentration on standards and results and the measurement of student achievements closely linked up to competition and accountability. Reward systems are tied to testing outcomes in an effort to change teaching practices in ways that increase student achievements. National testing systems influence strongly what goes on in classrooms, and this influence runs counter to the development of the learning and assessment culture that I have advocated.

Our Nordic countries have quite another assessment and learning tradition than the US and I do not fear a repetition of the California story in our countries. Nevertheless I think that we need to be alert in order to avoid that the testing culture, which has already got a stronghold in our political establishments, invades our classrooms and is allowed to destroy the emerging shift in learning culture. There are already signs in Norway on a new focus on testing. Again I will use *writing* as an example. Our present Minister of Education initiated in 2002 a process of developing tests in reading, writing and mathematics at different levels in school. The task given to the "test developers" was to design ICT-based tests which could be machine scored. This raises the very important question: What kinds of knowledge and skills and competences *can* be scored this way, and the even more crucial question: *What will be left out?* Then we are back to Lauren and Daniel Resnick's caveats about assessment policy:

1. *You get what you assess.* Educators will teach to tests if the tests matter in their own or their students' lives.
2. *You do not get what you do not assess.* What does not appear on tests tends to

disappear from classrooms in time.

3. *Build assessments toward which you want educators to teach.* Assessments must be designed so that when teachers do the natural thing - that is, prepare their students to perform well-they will exercise the kinds of abilities and develop the kinds of skills and knowledge that are the real goals of educational reform. (Resnick & Resnick p. 59)

The learning culture in writing in Norwegian classrooms have changed considerably over the last decades because of process oriented writing and improved assessment. The good result of this has been just documented in a large research study by the professors Berge, Evensen and Hertzberg (Evensen & Vagle 2003) as part of the Evaluation of Curriculum Plan 1997 for our basic school.

When the new test concept was introduced by the Ministry, the community of teachers and writing researchers in Norway feared that it would endanger what had been achieved in Norwegian writing classrooms and possibly destroy the new learning culture in writing, just as the California STAR test did. Therefore the three above mentioned wrote a letter to the Minister of Education, citing the substantial research findings both nationally and internationally which speak against this kind of reductive testing of writing.

I find this a good, although rare, example of how researchers use their expertise to influence the political processes when it becomes obvious that political decisions counteract the knowledge base of the field. In this particular case the professors were invited in on the process and have taken on the responsibility of developing tests worth teaching to. We still await the result. But the change in the educational climate in Norwegian schools due to the introduction of accountability and high stakes testing can already be noticed when talking to teachers and parents of children in school, and I have been told that the same thing has happened here in Iceland.

### 3. Some challenges for research

If the new modes of assessment are going to survive, research is needed both in order to improve the quality of the assessment practices and to investigate their consequences. High claims are being made at the moment about their positive influence on learning. Research is needed:

- to clarify theoretical basis of new modes of assess
- to document learning effects and critical factors
- to understand how to improve feedback and use
- to investigate constructive use of criteria
- to develop self and peer assessment
- to expand traditional concepts of validity & reliability (from psychometrics to edumetrics)
- to improve the quality of new modes of assessment

### Conclusion

A main message in this lecture has been that assessment is such an integral part of the teaching and learning culture that we cannot change one without the other. I have also reminded us all that assessment has always been a political issue, and that we as educators and researchers should not be



naïve or complacent about what drives the political agenda when it comes to education. In this day and age accountability, effectivity and quality management are key concepts in most countries. There is a strong tendency to think and act as if quality equals what can be measured, and the backlash in assessment policies that we have witnessed take place in the US, should be a warning. National tests vary in quality, and I have shown the importance of creating tests worth teaching to, as well as the importance of alternative forms of assessment. Our countries are small and I believe as educational researchers we may influence decision-making processes. But to be taken seriously we need a solid research base, and this is perhaps the biggest challenge to all of us.

## References

- Allern, M. (2003) Med en fot i det gamle? Hvilke konsekvenser får mappevurdering for lærer- og studentrollen? I Dysthe, O. & Engelsen, K. S. (red.) (2003) *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. Oslo: Abstrakt forlag.
- Biggs, J. (1996) Enhancing teaching through constructive alignment. *Higher Education* 32, pp 347-364.
- Birenbaum, M. (1996) Assessment 2000: Towards a pluralistic approach to assessment. In M. Birenbaum & F. J. R. C. Dochy (eds) *Alternatives in assessment of achievement, learning processes and prior knowledge* (pp. 3-29) Boston, MA: Kluwer.
- Black, I, mfl. (red.) (1994) *New directions in portfolio assessment. Reflective practice, critical theory and large-scale scoring*. Portsmouth, NH: Boynton/Cook, Heinemann.
- Broadfoot, P. (1994) *Educational assessment: The myth of measurement*. Bristol, Eng.: University of Bristol.
- Brown, G. Bull, J. & Pendlebury (1997) *Assessing student learning in higher education*. London/New York, Routledge.
- California Education Roundtable (1988) Systemwide and statewide assessment in California: A report of the subcommittee on student assessment to the California education roundtable. Sacramento, CA: Author.
- Clerehan, R., Chanock, K. More, T. & Prince, (2003) A testing issue: key skills assessment in Australia. *Teaching in Higher Education*, Vol. 8, 2, pp 279-284.
- Cooper, C. & Murphy, S. (1989) What we know about teachers' perception of the CAP and its' impact on instruction. Paper presented at the California Assessment Program "Beyond the Bubble" Conference, Anaheim, CA.
- Darling-Hammond, L.(1989) *Accountability for professional practice*. Teachers College Record, 91 (1), 60-80).
- Dewey, J. (1933) *How we think: A restatement of the relation of reflective thinking to the educative process*. Lexington, Mass.: Heath.
- Davis, A. & LeMahieur, P. (2003) Assessment for learning: Reconsidering portfolios and research evidence. . In Segers, M. Dochy, F. & Cascallar, E. (eds.) *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht/Boston/London, Kluwer Academic Press.
- Dierick, S. & Dochy, F. (2001) New lines in edumetrics: new forms of assessment lead to new assessment criteria. *Studies in Educational Evaluation*, 27, pp 307-329.

- Dochy, F., Seegers, M. & Sluijsmans, D. (1999) the use of self-, peer and co-assessment in higher education: a review. *Studies in Higher Education*, 331-350.
- Dysthe, O. et al. (1996) "Jeg vet hva jeg kan bra, hva jeg kan bedre og hva jeg kan mindre bra. Forsøk med mappevurdering i norskfaget i ungdomsskolen". I K. Skagen (red): *Karakterboka*. Oslo: Universitetsforlaget.
- Dysthe, O. (1999) «Mappevurdering som læringsform og sammenhengen med kunnskaps- og lærings syn.» I O. L. Fuglestad, S. Lillejord & J. Tobiassen (red.) *Reformperspektiv på skole- og elevvurdering*, Fagbokforlaget.
- Dysthe, O. (2001). *Dialog, samspel og læring*. Oslo: Abstrakt forlag.
- Dysthe, O. (2002) Professors as mediators of disciplinary text cultures. *Written Communication*. 19, 4, s. 458-493.
- Dysthe, O. & Engelsen, K. S. (red.) (2003) *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. Oslo: Abstrakt forlag.
- Dysthe, O. & Tolo, A.: Digital portfolios and feedback practices in a traditional university course. Paper presented at EARLI 10<sup>th</sup> Biennial Conference, Padova, Italy, Aug. 26.-30, 2003.
- Dysthe, O. & Engelsen, K. S. (2004) Digital portfolios as a learning tool in teacher education. *Assessment in Higher Education*, 29, No 2., pp
- Ellmin, R. & Ellmin, B. (2003). *Att arbeta med portfolio - teori, förhållningssätt och praktik*. Stockholm: Gothia.
- Evensen, L. S. & Vagle, W. (2003) *Kvalitetssikring av læringsutbyttet i norsk skriftlig. Sammendragsrapport for KAL-prosjektet*. Trondheim: Institutt for språk- og kommunikasjonsstudier, NTNU.
- Falchikov, N. & Goldfinch, J (2000) Student peer assessment in higher education. a meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70 (3) pp. 287-322.
- Gardner, H. (1984) *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Gielen, S., Dochy, F. & Dierick, S. (2003) Evaluating the consequential validity of new modes of assessment. In Segers, M. Dochy, F. & Cascallar, E. (eds.) *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht/Boston/London, Kluwer Academic Press.
- Gibbs, G., Simpson, C & Macdonald, R. (2003) Improving student learning through changing assessment - a conceptual and practical framework. Paper at EARLI Conference, Padova, Italy Aug 2003.
- Gipps, C. (1994) *Beyond testing. Towards a theory of educational assessment*. London: Falmer Press.
- Gillespie, C, Ford, K., Gillespie, R, & Leavell, A. (1996) Portfolio assessment: Some questions, some answers, some recommendations. *Journal of Adolescent & Adult Literacy*, 39, pp480-91.
- Greeno, J. Collins, A., Resnick, L. (1996) *Cognition and Learning*. D. Berliner & R. Calfee, (red.) *Handbook of Educational Psychology*, s. 15-46.

- Hamp-Lyons & Condon (2000) *Assessing the portfolio: Principles for practice, theory, research*, NJ: Hampton Press.
- Hauge, T. E. (2003) Læringsmapper som fornyingsredskap i praktisk-pedagogisk utdanning. I Dysthe, O. & Engelsen, K. S. (red.) (2003) *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. Oslo: Abstrakt forlag, pp. 183-2001.
- Hauge, T. E. & Wittek, L. (2002) Portfolios as mediators for collaborative learning and professional development in a distributed environment of teacher education. Paper to EARLI Assessment Conference: Learning
- Havnes, A., Taasen, I., Jensen, K. T., Lauvås, P. (2003) Mappevurdering som lærings- og vurderingsform. Erfaringer fra syskepleierutdanning. I Dysthe, O. & Engelsen, K. S. (red.) (2003) *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. Oslo: Abstrakt forlag.
- Hillocks, G. (2002) *The testing trap. How state writing assessments control learning*. New York: Teachers College Press.
- Immerwaht, J. (2000) Great expectations: How Californians view Higher Education. San Jose, C.: National Center for Public Policy and Higher Education and Public Agenda: Table 3, National column.
- Juul Jensen, M. (2000). "Man lærer at se sig selv og sine opgaver i øjnene". Om mappevurdering som alternativ vurderingsform. IN K: Esman, A. Rasmussen, L. B. Wiese (red): *Dansk i dialog*. København: Dansk lærerforening., pp s 91-120.
- Juul Jensen, M. (2003) Mappevurdering som alternativ vurderingsform. Afsluttende rapport om 2-årigt forsøg. Undervisningsministeriet. Prosjektnummer 2000-2434-0391.
- Kankaanranta, M. & Linnakylä, P. (1999). Digital portfolios as a means to share and develop expertise. *Portfoliota verkossa. Portfolios on the web*. P. Linnakylä, M. Kankaanranta, & J. Bopry (red.) s. 210-242
- Kankaanranta, M (2002). Developing digital portfolios for childhood education.
- Klenowski, V. (2002) *Developing portfolios for learning and assessment. Processes and principles*. London/New York: Routledge/Falmer.
- Krogh, E. & Juul Jensen, M. (2003) Portfolioevaluering - rapport om en ny evaluerings- og arbeidsform. Hæfte nr. 46. København: Utdanningsstyrelsen.
- Lauvås, P. & Jakobsen, A. (2002) *Exit eksamen eller? Former for summativ evaluering i høgre utdanning*. Oslo: Cappelen Akademisk Forlag.
- Lave, J. & Wenger, E. (1991) *Situated learning. Legitimate peripheral participation*. New York: Cambridge University Press.
- Lewin, K. (1952) Group decision and social change. In G. E. Swanson et. al (eds) *Readings in social psychology*. New York: Holt.
- Lindström, Lars (1992). Processinriktad utvärdering och bedömning. *Forskning om utbildning*, no. 3, pp. 34-44.

- Lindström, Lars (1994). Portföljmetodik i estetiska ämnen. Paper framlagt vid Nordisk Förening för Pedagogisk Forskning Kongress i Vasa 10-13 mars, 1994.
- Lindström, Lars (1997). *Portföljen som examinationsform*. Lärarhögskolan i Stockholm: Didaktikcentrum (Häften för Didaktiska Studier 64).
- Lindström, Lars (1999). *Portföljvärdering av elevers skapande i bild*. With Leif Ulriksson and Catharina Elsner. Skolverket / Liber. Rapport inom Skolverkets projekt "Utvärdering av skolan 1998". Incl. CD-ROM and videotape.
- Lindström, Lars (2002). Produkt- och processvärdering i skapande verksamhet. In: Skolverket *Att bedöma eller döma. Tio artiklar om bedömning och betygssättning* (pp. 109-124). Stockholm: Skolverket / Liber Distribution.
- Lindström, Lars (2003). Creativity: What is it? Can teachers assess it? Can it be taught? In: W. Rogala & S. Selander (Eds.) *Technology as a Challenge for School Curricula* (pp. 114-131). Stockholm: HLS Förlag, 2003.
- Lindström, Lars (2004). Novice or expert? Conceptions of craftsmanship among teacher educators and artisans. Paper to be presented at the Second Biannual Joint Northumbria/EARLI SIG Assessment Conference, University of Bergen, June 23-25.
- Ludvigsen, S. & Flo, C. (2002) Innovasjon i lærerutdanningen: Hvordan skapes endring. I S. Ludvigsen & c. Flo: *Et utdanningssystem i endring. IKT og læring*. Oslo: Gyldendal Akademiske Forlag.
- Ludvigsen, S. & Hoel, T.L (2002) *Et utdanningssystem i endring*. Oslo, Cappelen Akademisk Forlag
- McDowell, L. & Sambell, K. (1999) The experience of innovative assessment: Student perspectives. In S. Brown & A. Glasner (eds.) *Assessment matters in higher education*. The society for Research into Higher Education & Open University Press.
- McNeil, L. (2000) *Contradictions of school reform. Educational cost of standardized testing*. New York: Routledge.
- Molloy, G. (1996). *Reflekterande läsning och skrivning*, Lund: Studentlitteratur.
- Murphy, S. & Underwood, T (2000) *Portfolio practices: Lessons from schools, districts, and states*. Norwood, Mass: Christopher-Gordon Publishers.
- Murphy, S. (2003) That was then, this is now: The impact of changing assessment policies on teachers and the teaching of writing in California. *Journal of Writing Assessment*, Vol 1 No 1. pp. 23-45. Hampton Press Inc.
- Myers, M (1996) Saling ships: A framework for portfolios in formative and summative systems. IN R: Calfeed & P. Perfumo (Eds) *Writing portfolios in the classroom. Policy and practice, promise and peril* (pp 149-178). Mahwah, N. J.: Lawrence Erlbaum Ass.
- Otnes, H. *Collective digital text development in teacher education. Hypertext and dialogism in digital portfolios*. Paper at NERA conference, Tallin, 7.-10.mars  
<http://ikt.hive.no/ans/ho/Hypertextual%20writing.htm>
- Otnes, H. (2002) Skjermbaserte fellestekster. Digitale mapper og hypertekstskrivning. I: Hoel, T.

- Løkensgard & Ludvigsen, S. (red.): *Et utdanningssystem i endring*. Oslo: Gyldendal.
- Otnes, H. (2003) Læringsarena eller arkivskuff? Lærings- og dokumentasjonssjangrer i digitale mapper. I Dysthe, Olga & K. S. Engelsen (red.): *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. Abstrakt forlag.
- Packer, M. J. & Goicoechea, J. (2000) Sociocultural and constructivist theories of learning: ontology, not just epistemology. *Educational Psychologist*, 35 (4) , pp 227-242.
- Parliamentary proposition 27/(2000-2001): *Gjør din plikt - Krev din rett. Kvalitetsreform av høyere utdanning*. (<http://www.odin.dep.no/ufd/norsk/publ/stmeld/014001-040004/index-dok000-b-n-a.html>) (Do your duty - demand your rights. Quality Reform of higher education)
- Paulson, F. L., Paulson, P.R. & Meyer, C. (1991) What makes a portfolio a portfolio? in *Educational Leadership* 48, 5, p. 60-63.
- Pearson, P.D. (1987) Assessment, accountability, and professional prerogative. In J. E. Readance & S. Baldwin (Eds.) *Research in literacy: merging perspectives*. Thirty-sixth Yearbook of the National Reading Conference (pp. 3-16) Rochester, NY: National Reading Conference.
- Raaheim, A. (2000) Learning, knowledge and assessment at university. In V. Fahlen, A.M. Luhanen, L. Peterson & B. Stensaker (eds) *Towards best practice. Quality improvement initiatives in Nordic Higher Education*. Copenhagen: Temanord, Nordic Council of Ministers 2000: 501.
- Raaheim, A. & Raaheim, K. (2003) *Eksamen - en akademisk hodepine*. Bergen: Sigma.
- Resnick, L. & Resnick, D. (1992) Assessing the thinking curriculum: New tools for educational reform. In B. Gifford & M. C. O'Connor (Eds.) *Changing assessments: Alternative views of aptitude, achievement and instruction*. Boston: Kluwer Academic.
- Segers, M. Dochy, F. & Cascallar, E. (eds.) (2003) *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht/Boston/London: Kluwer Academic Press.
- Segers, M. Dochy, F. & Cascallar (2003) The era of assessment engineering: Changing perspectives on teaching and learning and the role of new modes of assessment. In Segers, M. Dochy, F. & Cascallar, E. (eds.) *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht/Boston/London, Kluwer Academic Press.
- Shavelson, R. J. & Huang, L. (2003) Responding responsibly to the frenzy to assess learning in higher education. *Change* Jan./Feb., pp11-19.
- Shepard, L. A. (2000) The role of assessment in a learning culture. *Educational Researcher*, Vol. 29, 7, s. 4-14.
- Shepard, L. A. (2001) The role of classroom assessment in teaching and learning. I: Richardson, V. (red.): *Handbook of research on teaching, 4<sup>th</sup> edition*. Washington, DC: American Educational Research Association.
- Sluijsmans, D. (2002): *Student involvement in assessment. The training of peer assessment skills*. Dr. dissertation, Open University, Netherland.
- Sluijsmans, D., Brand-Gruwel, S. & Van Meerienboer, G. (2002) Peer assessment training in teacher education: effects on performance and perceptions. *Assessment & Evaluation in Higher Education*.

Vol 27 (5), 443-453.

Säljö, R. (2001) *Läring i praksis*. Oslo: Cappelen Akademisk Forlag.

Takle, E. S, Sorensen, E. K., Taber, M. R., Fils, D..(2000) Structuring the past, present and future through virtual portfolios. In L. Dickinck\_Holmfeld & B. Fibiger (eds.). *Learning in virtual environments*. Copenhagen: Forlaget Samfundslitteratur

Topping, K. (2003) Self and peer assessment in school and university. In Segers, M. Dochy, F. & Cascallar, E. (eds.) (2003) *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht/Boston/London, Kluwer Academic Press.

Wenger, E. (1998) *Communities of practice. Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Wenger, E, McDermott, R, & Snyder, W. M. (2002) *Cultivating communities of practice*. Boston, Ma: Harvard Business School Press

Wittek, L. (2002) *Fra vurderingsinstitusjon til læringsinstitusjon? En undersøkelse av mapper som vurderings- og undervisningsform ved Universitetet i Oslo*.

Wittek, L. (2003) *Læringsarbeid og mappevurdering. En casestudie fra to profesjonsstudier ved det utdanningsvitenskapelige fakultetet*. Universitetet i Oslo.

Wittek, L. & Dysthe, O. (2003) Kvalitetskriterier og vurdering - et utviklingsområde innenfor høyere utdanning. I Dysthe, O. & Engelsen, K. S. (red.) (2003) *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. Oslo: Abstrakt forlag.

“...when fully implemented, CAP planned to assess eight different types of writing at Grade 8: a proposal for solving a problem, a sketch of a person whom the writer knows, a story, an evaluation (e.g. a judgement of a literary text, book, movie, artwork, etc) an autobiographical incident, observational writing, and an essay that speculates about causes or effects (Ibid. p 25). At Grade 10 the last two writing types would be assessed, namely an essay about a controversial issue, an interpretation of a literary work and a reflective essay. Each student's writing was to be evaluated on “Criteria emphasizing the critical thinking, problem-solving, and composing requirements unique to the particular type of writing it represented” (California Education Roundtable, 1988, p.52).

The program is called “Tuning education strategies in Europe”.

In Scandinavia the testing tradition has not been as strong as in the US, and the “alternative assessment” is more often seen as alternative to traditional individual pen-an-pencil school exams at the end of term.

‘Authenticity’: “to make the boundaries between school and the world more porous by bringing authentic contexts into classrooms, and at the same time, by developing habits of inquiry in schools that will make students good thinkers and problem solvers in the word” (Resnick 1987). This line of thought was anticipated by Dewey.

The shift of purpose from assessment for selection to assessment to support student learning comes from an understanding that “Learning is not possible without thoughtful use of quality assessment information by the learners. This is reflected in Dewey's (1993) “learning loop”, Lewin's (1952) “reflective feedback” and Wiggin's (1993) “feedback loop” ....If students are to develop into life long, independent, self-directed learners they need to be included in the assessment process so the “learning loop” is complete. Reflection and assessment are essential for learning. In this respect, the

concept of assessment for learning as opposed to assessment of learning has emerged” (Davis & LeMahieu 2003, p142).

“Variety in assessment techniques is a virtue not just because different learning goals are amenable to assessment by different devices, but also because the mode of assessment interacts in complex ways with the very nature of what is being assessed “(Shepard 2001:1087).

To put it very simply, hereditarian intelligence theory means that children are born intelligent or dumb, and IQ tests or objective tests would measure what they could achieve. The social efficiency curriculum aimed at eliminating waste, for instance the waste of teaching students anything they would not need in their future work, thus academic subjects should be taught only to students going to college. Even though such theories lost their popularity in the US in the 1930ies, they continued to influence curriculum thinking, according to Shepard. Behaviorist learning theories supported the teaching and testing of atomistic pieces of knowledge and there existed an alignment or congruence between dominant learning and curriculum theories underpinning teaching practices and standardized tests.

### **The European Language Portfolio**

The European Language Portfolio is a documentation system for language proficiency based on the Common European Framework of Reference, designed as a pedagogical instrument to carry through the intensions of the Framework. It represents a top-down assessment initiative with considerable consequences for the learning culture in each classroom. Since 18 European countries already are involved, including the Nordic countries, it is indeed very influential. The portfolio contains detailed descriptions of levels of competence for listening, reading, spoken interaction, spoken production and writing in the foreign language. This Portfolio has the potential of changing foreign language teaching in the classroom and beyond, and it is envisioned as an instrument to empower the students to become autonomous learners and take responsibility for goal setting and self assessment. But importing it wholesale may do more harm than good: professional development of teachers is needed, and experience has shown that it takes time because it involves contextualising and grounding the practical pedagogy in views of learning that are compatible with the overall values of the national school system. Finland, for instance, has taken up this challenge and involved several universities in their 3 year European Language Portfolio project, which was led by Viljo Kohonen and Ulla Pajukanta, Tampere University.

<http://www.jyu.fi/agora-center/dpf/research.shtml>

<http://www.jyu.fi/agora-center/all/kontaktie.html>

<http://www.jyu.fi/ktl/publications/electronic/001/kankaanranta.htm>

The project is documented in the book Dysthe & Engelsen (eds.): *Mapper som pedagogisk redskap [Portfolios as pedagogical tools]* Abstrakt forlag 2003.

(<http://stud.no/prosjekt/nordbygd>)

Ludvigsen & Flo (2001) in an activity theory based analysis of portfolios in teacher education at the University of Oslo, called the portfolio 'a boundary object' because it established dialogues across learning sites and across subjects as well as integrated theory with practice.

See McDowell & Sambell (1999) “The experience of innovative assessment: Student perspectives”.

There exist a substantial international research literature on peer and self assessment, and even review articles of such literature i.e. Dochy, Seger & Sluijsmans 1999, Falchikow & Goldfinch 2000,

Topping 1998, 2003.

10

Shepard, 2001, p. 1067

Emergent Paradigm  
(ca 1990-2000+)

Dissolution of Old Paradigm: New Views of Instruction/Old Views of Testing  
(ca 1980-2000+)

20<sup>th</sup> cent Dominant Paradigm (circa 1900s -2000+)

Cognitive & Constructivist  
Learning Theories

Reformed Vision of Curriculum

Classroom Assessment

Constructivist Curriculum  
&  
Instruction

Traditonal Testing  
Scientific Measurement

Social Efficiency Curriculum

Hereditarian  
Theory of IQ

Associationis  
& Behaviorist

Learning

Theories

Scientific Measurement