

PROBLEMS INDUCED BY AMALGAMATION OF PEDAGOGICAL PROGRESSIVISM AND EDUCATIONAL ACCOUNTABILITY: ORAL EXAMS WITH PRIOR PREPARATION TIME IN NORWEGIAN SECONDARY SCHOOLS

Eyvind Elstad, Are Turmo

University of Oslo, Norway

E-mail: eyvind.elstad@ils.uio.no; are.turmo@ils.uio.no

Øystein Guttersrud

University of Oslo, Norway

E-mail: oystein.guttersrud@fys.uio.no

Abstract

The progressive movement in Norwegian education has had the hegemony among civil servants, politicians and major groups of teachers. However, during recent years the assessment system has moved towards increased educational accountability based on performance measures (exams and national tests). At the same time, progressivistic inspired assessment has remained: oral exams with prior preparation time. The pupils are typically given the topics 48 hours in advance and have the opportunity to prepare at home or at school. They may take the exam individually or in groups. The teacher, who also asks the questions during the exam, defines the topics. The grades are assigned by an external evaluator, but the subject teacher should also take part in the assessment. This amalgamation of pedagogical progressivism and educational accountability induces major challenges in education.

Key words: educational accountability, progressive school philosophy, oral exam, assessment.

Introduction

The progressive movements in education for centuries had the hegemony in Norwegian education among bureaucratic officeholders (Skagen, 2004), educational researchers and major groups of teachers, especially in primary schools. For instance, in July 2001, one of Norway's leading newspapers printed an article headed "Reform attempts in 700 schools" (Kluge, 2001). Reformation of teaching methods was discussed. One of the top civil servants in the Ministry of Education and Research, Director General Ole Briseid, condemned teacher-centred instruction and emphasized progressive, student-oriented and activity-oriented teaching methods:

The Ministry has wanted to reform the teaching methods in schools (in Norway) for a long time. Now this is finally happening. The Ministry has wanted this for a long time. He (Briseid) wants more project work, less traditional teaching using the blackboard as a visual aid and more problem-based teaching. To a larger extent, pupils will work independently and in groups with topics they develop themselves. The teachers will be more like supervisors than lecturers, and computer technology forces the development of new teaching methods (Kluge, 2001, p. 3).

This is an example of strong process steering after the millennium. This example illustrates how the education policy in Norway was aligned with ideas linked to a progressive

school philosophy (abbreviated PSP). Briseid's proclamation of steering work processes was in accordance with the Labour Party's governmental policy at the time (Skagen, 2001). The self-satisfaction with the education policy was high (despite Norway in 1995 having the lowest science-score of all European countries in Trends in International Mathematics and Science Study (TIMSS), Lie et al., 1997). Later in 2001, an educational county-director declared that, "The Norwegian school has never been better" (Skrede, 2004, p. 7).

However, a turnaround in educational policy and management was initiated by a centre-conservative government in 2002. The disappointing Norwegian results from the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment — PISA 2000 — gave legitimacy to this turnaround (OECD, 2010), and during the years that followed, the assessment system moved toward increased accountability based on performance measures. The advent of educational accountability (abbreviated EA) after the millennium has increased external pressure for improved performance, and performance measures are used as indicators of goal attainment. However, progressivistic ideas and the educational zeitgeist of accountability are now embedded side by side in Norwegian assessment policies for oral exams. This situation is referred to as the *amalgamation* of PSP and EA.

The purpose of this article is to discuss, interrogate and identify problems inherent in the tensions between the ideals of PSP and precise and adequate measurements of pupil performance in the service of EA. These tensions induced challenges and problems of Norwegian education in the beginning of the 21st century. The analytical framework used serves the purpose of seeking to understand the assessment practices as a consequence of diverging values, norms and preferences. **We constrain our study to the exploration of oral assessment policies and practices in Norwegian lower secondary education.** More specifically, we explore the extent to which grades from oral exams with prior preparation time may reasonably be used for accountability purposes, based on analysis of available grade statistics. To accomplish this we explore the national results statistics from the years 2002-9.

Background of the Amalgamation of Progressive School Philosophy and Educational Accountability

Norway was a latecomer in terms of applying accountability devices in education in the 1990s (Elstad, Nortvedt & Turmo, 2009). The OECD produced test and assessment policy recommendations for the Norwegian educational authorities in 1988 (OECD, 1988). Then, the government secretaries of state replied in the 1990s by adjusting the balance of power in the educational sector more in favour of central government (White Paper 37, 1990-91), using regulations to oblige schools to implement project work and imposing process-descriptive targets onto the elementary-school curriculum. The key attribute of progressivism is project work (W.W.C., 1922). In connection with the Norwegian school reform in 1997 (KUF, 1996), it became a legally binding regulation that at least 60 % of the time in lower primary school and at least 30 % in upper primary school should be spent on theme-based teaching and project work (Helgeland, 2001). This is an example of detailed control of schools' teaching methods (Tønnesen, 2004). The ideal was the pupil-centred school tasks emphasizing pupils' intrinsic motivation for learning. This philosophy emphasised the importance of giving pupils opportunities to be involved in the choice of teaching methods and topics. The pupils should also develop problems and seek relevant sources (Sejersted, 2005; Skagen, 2001). The proponents of this progressive school philosophy were hostile to use of external incentives such as grades and examinations (Telhaug et al., 2006). Grades and examinations may result in a sense of humiliation and stigma, that is to say, experiences that could have a negative effect on the pupils' desire to learn. A theoretical underpinning was the emergence of the concept intrinsic motivation (Deci, 1975) and how extrinsic motivation instigated by the teacher may crowd out

the pupil's intrinsic motivation.

During the first decades following the Second World War:

The Scandinavian countries were particularly oriented towards international reform/pedagogic theory, its appreciation of the child's personal potential and the desire to place the pupil at the centre (...) Scandinavian educational policy of the 1970s made a more explicit attempt than ever before to implement progressive, pupil-centred and activity-oriented teaching methods. (Telhaug et al., 2006, p. 254).

In the 1990s, school documents maintained that the goal was "to create equality between pupils and teachers in the democratic teaching organisation where the pupils set the terms and make contributions" (NOU, 2003, p. 103). Progressivism also manifested itself in the way learning objectives were expressed in the curriculum, and in the assessment criteria determined by the civil service, such as "pupils are to work with numbers", "learn to read at their own speed", and "work with some interesting numerical relations for instance numbers with special properties, the role of numerology in certain cultures or the attraction of number riddles" (Ministry of Education, 1996, p. 187). Assessment for learning was particularly emphasised in this reform (Ministry of Education, 1998). Here, the subject's core and distinctive characteristics are given less emphasis than the pupil's experience of something that is considered valuable by progressive pedagogues (Telhaug et al., 2006). **Another characteristic of the educational efforts in the 1990s, which can also be linked to progressive ideas, is the introduction of *the open-book concept in connection with examinations*, which is strongly linked to the topic of the present article.** This concept entails the pupil having access to all resources for some written examinations, as well as preparation time and the opportunity for external help to solve previously notified tasks for the oral exam. This idea originated from researchers (Eilertsen & Valdermo, 1997), and was tested in a pilot project, evaluated (Mathisen, 1997) and thereafter, introduced as a norm for assessment practice. Part of the intention has been to achieve a greater degree of authenticity in the examination situation. This idea must also be viewed in connection with developing ideas concerning authentic assessment (Wiggins, 1990) and situated cognition and the culture of learning (Brown, Collins & Duguid, 1989). Preparation time for oral exams and the opportunity to take useful resources into written exams are practices that have been in place for some time, although in modified forms. The above indicates that the Norwegian educational policies of the 1990s were different from the tendencies seen in many other European countries, for instance England and the Netherlands (Christophersen et al., 2010).

Shortly before the publication of the PISA 2000 results in December 2001, Norway attained a new centre-conservative government (in office 2001-2005), in which a representative from the Norwegian Conservative Party (Høyre) became the Minister of Education and Research. The PISA results gave her a flying start as a minister. Her Deputy Minister claimed that the publishing of the PISA results set the scene for what he referred to as "a battle over knowledge" in the Norwegian school system (Bergersen, 2006, p. 40-41) and gave the legitimacy of a systemic shift towards performance controlled by the use of measurements (exams and later national tests and school performance indicators). A novel educational discourse of result orientation in education was established in Norway (Elstad, Nortvedt & Turmo, 2009), with the advent of the new educational policy (assessment-based accountability) by the conservative-conservative government. **Norway was inspired politically by policy implementation in Great Britain in making this turn legitimate** (Christophersen et al., 2010).

Several studies suggest that assessment-based accountability policies may be effective at raising pupil attainment (Hanushek & Raymond, 2005). Accountability systems in education are designed to improve pupil attainment. The notion behind assessment-based accountability is that assessment will provide pupils, teachers, and heads with an incentive to work harder, as

well as help to identify struggling pupils and schools. In general, accountability systems have not been approved for use in the Norwegian nationwide education sector, but accountability thinking nevertheless influences how some of Norway's local quality-assurance systems are designed. Certain building blocks of an accountability system were, nevertheless, introduced in 2004 and later: performance controlled by the use of national tests and examinations, school-contribution indicators based on tests and examinations, publication of average results by school, among others. The Norwegian Parliament unanimously agreed to introduce testing, decentralize responsibility for quality of schooling to municipalities, auditing and supervision, and new curriculums with clearly defined goals emphasising basic reading, writing and arithmetic skills (Ministry of Education, 2006). Teachers were to give greater academic pressure in teaching (White Paper 30, 2003-4) as part of an ambition to raise the level of knowledge in Norwegian schools.

In Norway's 2005 general election, a red-green coalition won the majority. The political platform of the Ministers of Education and Research from Socialist Left Party (Sosialistisk venstreparti), initially opposed standardized national tests, and was in favour of removing grades and examinations from lower and upper secondary education. However, the red-green government has chosen to retain some of the features of educational accountability, but has weakened the potential for accountability *mechanisms*. The responsibility for quality in schools has been delegated to the governing bodies, but the extent to which counties and local authorities have been able to relate to this type of delegated responsibility has varied (White Paper 31, 2007-8). The red-green government has made it possible for governing bodies to develop local accountability systems (ibid). In a regime of assessment-based accountability, reliable, acceptable and precise measurements of learning, carried out under controlled conditions, will be a prerequisite for the appropriate functioning of the system.

Tests and exams that serve accountability purposes need to be reliable, valid and fair. To achieve these goals, the tests have to be standardized and objective. If stakes are not high, then other considerations (e.g. the examinee being relaxed) may be pertinent. A number of challenges are related to the use of imprecise school accountability measures (Kane & Staiger, 2002). For instance, prior pupil attainment can help increase the precision of measurements relating to the school's particular contribution to pupil progress. One problem related to the Norwegian school system is that there are so few reliable and valid measurements of pupil performance at different age levels that it is difficult to estimate value-added performance. It is, however, not difficult to agree with "the need for more rigorous evaluation of the implementation of (incentive) schemes in order that their design may be improved in a systematic way" (Propper & Wilson, 2003, p. 265). However, such tenets in an accountability system for education may be in conflict with PSP ideas about authenticity in the examination situation and that the pupil actually learns something whilst the exam is being carried out. In this way, tensions arise between deeply contradicting educational philosophies (PSP and EA) that both have positive, but different intentions. In an educational system based on making school officials responsible, these tensions will create challenges in terms of handling exam regulations. As mentioned, we restrict ourselves to discussing these tensions in connection with oral examinations in lower secondary education at year 10.

Oral Exams with Prior Preparation Time as an Incidence of Amalgamation

Powers and responsibilities have been devolved from national level to local levels. The school owners – the local educational authorities – are responsible for implementing oral exams in Norwegian secondary schools. Norway has a long tradition for school-leaving exams in both lower and upper secondary school (Lysne, 2004; Jarning, 2010). Beyond 1997, oral exams with prior preparation time have become common practice. Examinees are selected randomly for oral

testing. The examinees are typically given a specified topic 48 hours in advance of the exam. During this time, the pupils may prepare at school or at home. Examinees may take the exam in pairs or in larger groups. The rationale behind these exam procedures, as mentioned earlier, is strongly related to the notion of implementing authentic situations in the examination.

Oral examinations with preparation time may be organised in different ways. The pupil may, for instance, produce a Power-Point presentation with a 10-minute commentary. Thereafter, the teacher asks in-depth questions (about 15 minutes) on the given topic. The pupil then answers concrete questions covering another topic posed by the teacher (about 10 minutes). The external evaluator does not ask the examinee direct questions, but can request that the teacher ask specific questions. However, the external evaluator has the final word regarding the grade assigned. We have briefly described to examples from Norway below.

One concrete example is from a year-10 oral exam in mathematics, where the pupils were given 48 hours in which to imagine and present a mathematical story connected to a given topic, for instance “going on a journey”. The evaluation criteria include “the ability to show creativity in the form of the presentation”, “the ability to show interest and enthusiasm”, “the ability to co-operate” and “the ability to show respect for the opinions of others” (Hansen 2004, p.24; our translation). Another example includes oral exams in the school subject of Norwegian (Bakke & Kverndokken, 2010). “Developments in Norwegian language in the period 1830 to 1917” is an example of a topic. After the presentation, the pupils may be asked in-depth questions related to, for example, “realism in Norwegian literature”.

One may argue that these examination procedures are authentic, testing the pupils’ skills in something approaching real-world settings, e.g., their ability to collect and structure information from different sources, communicate a clear message. However, the reliability and comparability of these oral exam results may be questioned, for several reasons. Firstly, the teacher’s influence over the results may be rather strong, as the teacher, not the examiner defines the tasks. This is in contrast to national centralised written exams, where the content is unknown to teachers and pupils until the examination day. Secondly, the teacher is also involved in the assessment and the assigning of grades in the oral exams, even though the external examiner technically has the final word. Again, this stands in contrast to written exams, where the grades are anonymously assigned by external evaluators. Finally, the pupils may receive help from parents and/or other persons, and pupil access to cultural and social capital during the examination process may be an important issue.

Methodology of Research

To what extent may grades from oral exams with prior preparation time (an idea originated in PSP) reasonably be used for accountability purposes, based on the analysis of available grade statistics? The purpose of this investigation is to explore the problems that arise with the amalgamation of PSP and EA and to examine available grades statistics for year 10 /16-year-olds) to explore the following research questions:

1. How do oral exam grades compare to the corresponding teacher given grades and written exam grades, and how stable are these relationships over time?
2. In relative terms, do oral exams favour one gender or any specific socio-economic group?

The Education Act in Norway dictates that all local authorities have to implement oral exams at year 10. Some local authorities may decide to implement oral exams with prior preparation time, whilst others will not. It is not possible to distinguish between these groups of local authorities in the available grade statistics. At present, however, an oral exam with prior

preparation time is by far the most common approach at the year-10 level.

In 2006, a new curriculum was implemented in Norwegian primary and secondary schools, while its precursor was implemented in 1997. Official grade statistics for Norwegian lower-secondary schools are available from 2002 and onwards, and exams only apply at the end of year 10. All grade statistics at year 10 in the period 2002-2007 are based on a PSP- inspired curriculum from 1997. Classical statistical test theory (Crocker and Algina, 1986) was used in analysis of grade data. Data are available for all the subjects in which written exam apply: Mathematics, English and Norwegian. Mean grades and averages for boys and girls, as well as for students with different socio-economic status (SES), defined by parental education level (International Standard Coding of Education, ISCED; UNESCO, 2006) were estimated and compared. Differences between girls and boys (positive values in favour of girls) and between the highest and lowest ISCED categories (positive values in favour of the highest ISCED group) were focused.

Results of Research

Table 1 compares the average teacher-awarded grades with year-10 exam grades at the national level in the period 2002-2007. Table 1 shows that in all subjects, the mean grades from oral exams are higher than the other grade averages. Furthermore, the average written- exam grades are the lowest among the grade averages in all the subjects in which written exams apply. These tendencies are strongest in Mathematics.

Table 1 explicitly displays differences between mean grades between 2002 and 2007. In most instances, the results show a weak “inflation” in mean grade level from 2002 to 2007. In all subjects other than English, the oral exam results display the largest difference between the two years. Again, Mathematics shows the strongest tendency. In 2002, the difference between the oral and written exam grade averages was 0.57. In 2007, this difference had increased to 0.89.

The results regarding oral exams in Mathematics and Science are particularly interesting when compared to the international comparative achievement studies’ results at the lower-secondary level during this period (Martin et al., 2008; Mullis et al., 2008; OECD, 2007). While the national oral exam results in these subjects show the strongest progress among all subjects, the international test results show a strong decline in average performance. However, the written exam results in Mathematics (no written exams apply in Science) show the same trend as in the international studies.

Table 1. Mean grades in Norwegian lower-secondary schools (year 10) in the period 2002-2007 (source www.ssb.no, N≈60 000). Oral exam results highlighted. Diff. =difference between 2007 and 2002.

	2002	2003	2004	2005	2006	2007	Diff.
Mathematics, teacher, written	3.44	3.47	3.45	3.47	3.45	3.45	+ 0.01
Mathematics, exam, oral	3.84	3.89	3.98	4.06	4.04	4.04	+ 0.20
Mathematics, exam, written	3.27	3.26	3.22	3.12	3.11	3.15	- 0.12
English, teacher, written	3.70	3.73	3.73	3.78	3.77	3.78	+ 0.08
English, teacher, oral	3.96	3.99	4.01	4.03	4.01	3.99	+ 0.03
English, exam, oral	4.31	4.34	4.35	4.35	4.33	4.34	+ 0.03
English, exam, written	3.52	3.53	3.58	3.59	3.64	3.64	+ 0.12
Norwegian, teacher, written	3.79	3.83	3.85	3.86	3.88	3.87	+ 0.08
Norwegian, teacher, oral	3.95	4.03	4.05	4.06	4.06	4.04	+ 0.09
Norwegian, exam, oral	4.23	4.30	4.32	4.36	4.38	4.36	+ 0.13
Norwegian, exam, written	3.60	3.60	3.67	3.64	3.65	3.66	+ 0.06

Subjects in which written exam does not apply:							
Social studies, teacher	3.98	4.02	4.04	4.05	4.05	4.03	+ 0.07
Social Studies, oral exam	4.19	4.21	4.29	4.33	4.31	4.32	+ 0.13
Science, teacher	3.85	3.90	3.92	3.95	3.95	3.95	+ 0.10
Science, oral exam	4.07	4.15	4.25	4.28	4.28	4.29	+ 0.22
Religion and ethical studies, teacher	3.93	3.96	3.99	4.01	4.01	3.99	+ 0.06
Religion and ethical studies, oral exam	4.21	4.23	4.29	4.36	4.34	4.36	+ 0.15

Table 2 shows mean grade averages for boys and girls, as well as for students with different socio-economic status (SES), defined by parental education level (International Standard Coding of Education, ISCED; UNESCO, 2006). Differences between girls and boys (positive values in favour of girls) and between the highest and lowest ISCED categories (positive values in favour of the highest ISCED group) are explicitly displayed. Results are available for the subjects in which written exam apply: Mathematics, English and Norwegian.

Further, Table 2 shows differences in favour of girls for all grade types, the differences in the subject Norwegian being the largest. In Mathematics, the largest difference in favour of girls is found for oral exams and the smallest for the teacher-allocated grades. In English, the difference for oral exams corresponds to the difference for written exams and teacher-allocated oral grades. In the subject of Norwegian, the gender difference is the largest for oral exams and teacher-allocated written grades. The difference between girls and boys for written exams is smaller. In summary, in both Mathematics and Norwegian, the oral exam format seems to favour girls relative to boys, when compared with the written exam format.

Table 2. Mean grades in Norwegian lower-secondary schools 2009 by gender and student socio-economic status, source www.ssb.no (N≈60 000).

	Parental education								
	All	Boys	Girls	Diff. gender	ISCED 0-2	ISCED 3	ISCED 5, short	ISCED 5, long	Diff. (high-low)
Mathematics, teacher, written	3.5	3.4	3.5	0.1	2.7	3.2	3.8	4.3	1.6
Mathematics, exam, oral	4.1	3.9	4.2	0.3	3.3	3.8	4.3	4.8	1.5
Mathematics, exam, written	3.4	3.3	3.5	0.2	2.6	3.2	3.8	4.3	1.7
English, teacher, written	3.8	3.6	4.0	0.4	3.2	3.6	4.1	4.5	1.3
English, teacher, oral	4.0	3.9	4.2	0.3	3.4	3.8	4.3	4.7	1.3
English, exam, oral	4.3	4.2	4.5	0.3	3.8	4.1	4.6	5.0	1.2
English, exam, written	3.8	3.6	3.9	0.3	3.1	3.6	4.1	4.4	1.3
Norwegian, teacher, written	3.8	3.5	4.1	0.6	3.2	3.7	4.1	4.4	1.2
Norwegian, teacher, oral	4.1	3.8	4.3	0.5	3.4	3.9	4.3	4.7	1.3
Norwegian, exam, oral	4.4	4.1	4.7	0.6	3.8	4.2	4.7	5.0	1.2
Norwegian, exam, written	3.4	3.2	3.6	0.4	2.9	3.2	3.7	3.9	1.0

Table 2 indicates that the difference between the extreme SES groups is the smallest in Mathematics for oral exams and the largest for written exams. In other words, in this subject,

the oral exam format seems to be relatively more favourable to the low SES group. The same tendency is established for the subject of English, although the tendency is weaker. In Norwegian, however, the smallest difference is established for written exams. It has to be emphasised that the contents of the written and oral exams in Norwegian are rather different.

Discussion

The main purpose of this research was to study the extent to which grades from oral exams with prior preparation time (which originated in SPS) may reasonably be used for accountability purposes, based on the analysis of available grade statistics. To accomplish this, the national results statistics were explored, which conveniently enough are available from 2002. The analysis shows a number of interesting results, which will be discussed below.

Firstly, attainment goals in the subject curricula are identical in terms of written and oral examinations (except for the subject Norwegian). Given the way in which goal statements and the school assessment guidelines are designed, the question of written or oral examination should be of secondary importance. This is definitely not the case, according to our findings. The analysis shows that the exam format itself has clear significance for the measured results. This phenomenon is regarded as an educational problem in that the grades achieved by the pupil should provide information about the attainment of targets in the goals laid out in the curriculum. When the type of assessment is significant for the result achieved, random elements will affect the results' statistics. For instance, a pupil with a given teacher allocated grade that is randomly selected for an oral examination in Mathematics has a greater chance of achieving a high grade than an equivalent pupil who is randomly selected for a written examination in Mathematics. This contributes to an element of unfairness as pupils' achieved grades have consequences for whether they match upper secondary schools' entrance levels. "Fairness" has particular significance in those areas of Norway where students compete for approval to attend certain schools.

It is – as shown in this study - generally advantageous to be selected for an oral exam, whilst having written exams in subjects such as Mathematics will tend to lower average grades, something that will be disadvantageous to the individual pupil who is in that position. Mathematics is the subject with the greatest problem in this respect, whilst English appears to be an anomaly in our material. In other words, this points to a weakness in the Norwegian examination system's tendency to be subject-dependent.

Secondly, when the above-mentioned weaknesses come into effect, the lower secondary schools' average grades in relation a certain subject become imprecise as a measure of the pupil's actual achievement. This weakness has particular consequences for the management logic created by results-based systems. Norwegian local authorities have been criticised by national bodies because the quality-assurance systems in their educational sector are inadequate (OAGN, 2006). Consequently, the systems have been improved. The weakness that is seen in the measured management data introduces the possibility that those involved may game the system. What we are indicating here can be seen as problems with imprecise measurements of pupil performance from a management perspective that places emphasis on accountability (Propper & Wilson, 2003). When schools and teachers are held responsible for their contributions to pupils' learning progress (Christophersen et al., 2010), any element of randomness and imprecision weakens the value of the grades statistics as management information.

Oral exams are neither objective nor standardized; they suffer from biases related to the examinee, the examiner, the interaction between them and the exam itself. In spite of their limitations, they might be appropriate for specific populations in some cases (e.g. examinees with learning disabilities) and certainly may be used for internal assessment by teachers, as they may involve the assessment of skills that are not measured by written tests.

The system of preparation time for oral exams in which parts of the examination content are known in advance may contribute to reinforcing social differences, as pupils can receive help during the preparation period. The data analysed does not give direct empirical support for this hypothesis, but it cannot be simply swept aside. Further research should be able to determine whether this system has a discriminatory effect.

On the other hand, those who defend the present system for oral exams emphasise the advantages in the pupil being able to learn whilst preparing for an exam (Dobson, 2009). This assumption is grounded in a progressivistic educational philosophy. Advocates will also emphasise that when the pupil has access to helpful resources whilst preparing, the situation will resemble authentic learning situations outside the school gates (Wiggins, 1990). When pupils use the preparation time well, they may actually gain knowledge while preparing for the oral exam. This promotes the schools' most important purpose – the pupils' learning (Ercikan, 2006).

In the case of written exams, independent experts develop the tests and students respond to them. In the case of the oral exams, it is the pupils' own subject teacher who devises the tasks. On one hand, a system like this may weaken the consistency in which oral exams are carried out. On the other hand, the system can be defended based on progressivistic ideas about examinations with a more personal approach. The teacher has another relation to the pupil besides being an external examiner. Through their knowledge of the pupils, the teacher can persuade them to perform well in a situation, which for many youngsters is full of nervousness and excitement. The sense of mastering the final school exam can contribute to motivation to continue in education, in a society in which completed education is highly significant in terms of subsequent employment opportunities. Drop-out from upper secondary school is a significant problem in Norway (Markussen, 2009), and where the transition from lower secondary to upper secondary schooling is based on a sense of mastery, this may help in maintaining the motivation to continue. There are a number of other progressivistic inspired justifications in favour of the current exam system.

The empirical results show that pupils with weak grades in Mathematics, combined with a low socio-economic background, form a group that benefits the most from the system of oral exams in Mathematics. A possible explanation is that this group of under-achievers is in a position to re-capture something of what they have lost when the oral exam presses them to exert effort over a limited period. In such a perspective, the actual amalgamation of PSP and EA orientations can be viewed as a useful means of promoting learning processes. On the other hand, this mechanism does not have validity for other subjects. We cannot escape the inference that the ideals of precise and adequate measurements of pupil performance in the service of educational accountability can stand in opposition to the visions of progressivistic pedagogy. Those who determine which systems should be used have the difficult task of weighing the advantages and disadvantages created by the different alternatives of the different options. As researchers, we have to be content with pointing out that our relationship to the tensions between different orientations, each of which represent legitimate intentions, is essentially a matter of values and priorities.

As with all similar studies, this study has certain limitations from a methodological as well as conceptual perspective. Firstly, the data discussed in this paper was only available to us at an aggregated level, i.e. it was not possible to compare different types of grades at the individual student level. Secondly, we had to use pre-defined socio-economic classifications made by Statistics Norway (the ISCED system), in which some of the ISCED categories were collapsed. More fine grained and detailed analysis of the effects of socio-economic status was not possible. Finally, the data available did not allow us to study potentially interesting interaction effects between gender and socio-economic status. We acknowledge these limitations and argue that they contribute to a foundation for future studies. This article serves as a starting point that will stimulate further research. More research is needed in order to come closer to inferences

about causality.

Conclusion

Despite its shortcomings, this study contributes to our initial understanding of the challenges of the amalgamation of EA and PSP. If the statistical inferences represent causal relationships, our findings may have implications for practice. The main conclusion is that the ideals of precise and adequate measurements of pupil performance in the service of educational accountability can stand in opposition to the visions of progressivistic orientation. According to EA, decision-making should be based on high quality evidence (Kane & Staiger, 2002). This evidence should include multiple items, formats (e.g. written tests, performance based), assessors and so forth. If one does not wish to get rid of the oral exam, it could serve as one component of an assessment system with pre-determined weights for each of its components.

References

- Brown, J. S., Collins, A., Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18 (1), 32-42.
- Bakke, J. O., Kverndokken, K. (2010). Å vurdere muntlighet – fra et norskfaglig ståsted [Assessing oral skills - a Norwegian language perspective] in Dobson, S. and Engh, R. (eds.) *Vurdering for læring i fag* (Assessment for learning in school subjects). Kristiansand: Høyskoleforlaget.
- Bergersen, O. H. (2006). *Kampen om kunnskapsskolen* (The fight for a knowledge based school). Oslo, Norway: Universitetsforlaget.
- Christophersen, K. A., Elstad, E., Turmo, A. (2010). Is teacher accountability possible? The Case of Norwegian High School Science, *Scandinavian Journal of Educational Research*, 54 (5), 413-429.
- Crocker, L., Algina, J. (1996). *Introduction to classical and modern test theory*. Fort Worth: Harcourt Brace Jovanovich College Publishers.
- Deci, E. (1975). *Intrinsic motivation*. New York: Plenum Publishing Co.
- Dobson, S. (2009). Muntlig vurdering – utfordringer og muligheter (Assessing oral skills- challenges and possibilities), in Dobson, S., Eggen, A. and Smith, K. (eds.) *Vurdering, prinsipper og praksis – nye perspektiver på elev- og lærlingvurdering* (Assessment, principles and practice - new perspectives on assessing students). Oslo: Gyldendal.
- Eilertsen, T. V., Valdermo, O. (1997). *Åpenbokvurdering: bidrag til bedre læring?* (Open book assessment: a contribution to better learning?). UNIKOM's Report no. 1 Tromsø: University of Tromsø.
- Elstad, E., Nortvedt, G. A., Turmo, A. (2009). The Norwegian Assessment System: An Accountability Perspective. *Cadmo*, 17 (1), 89-103.
- Ercikan, K. (2006). Developments in assessment of student learning, in Alexander, P.A. & Winne, P.H. (eds.) *Handbook of educational psychology* (2nd ed.) (pp. 929-952). Lawrence Erlbaum, Associates, Mahwah, New Jersey
- Hansen, R. G. (2004). *Bør økonomer kunne derivere?* (Do economists need to know how to differentiate?). *Samfunnsøkonomen*, No. 4/5.
- Helgeland, G. (2001). *Opplæringslova. Lov om grunnskolen og den vidaregående opplæringa av*

17. juli 1998 nr. 61. *Kommentarutgåve*. (The Norwegian School Act: Version with comments). Oslo: Universitetsforlaget.

Hanushek, E. A., Raymond, M. E. (2005). Does school accountability lead to improved performance? *Journal of Policy Analysis and Management*, 24 (2), 297-327.

Jarning, H. (2010). Resultater som teller. – Kunnskapskontroll og resultater i skolens århundre (Results that counts. Controlling knowledge and results in the school's century)). In E. Elstad & K. Sivesind (Eds.) *PISA – sannheten om skolen? (PISA - the truth about schools?)*. Oslo: Universitetsforlaget.

Kane, T. J., Staiger, D. O. (2002), The Promise and Pitfalls of Using Imprecise School Accountability Measures. *Journal of Economic Perspectives* 16 (4), pp. 91-114.

Kluge, L. (2001). Reformforsøk i 700 skoler (Reform trials in 700 schools), *Aftenposten* 5.7.2001, 4.

KUF (1996). *Læreplanverket for den 10-årige grunnskolen* (Curriculum for Primary, Lower and Upper Secondary Education). Oslo: Kirke, utdannings- og forskningsdepartementet. (R94 and L97). Available at: <http://www.undanningsdirektoratet.no/templates/>

Lie, S., Kjærnsli, M., Brekke, G. (1997). *9-åringers kunnskaper og holdninger i realfag i et internasjonalt perspektiv* (9 year olds' knowledge and attitudes in science and mathematics seen from an international perspective).TIMSS report no 25. Oslo: Department of Teacher Education and School Research, University of Oslo.

Lysne, A. (2004). *Karakterer og kompetanse. Kampen om skole. 1970-2000* Bind 2 (Grades and competencies. The fight over school) Oslo: AVA.

Markussen, E. (2009). *Videregående opplæring nesten for alle* (High school for almost all). Oslo: Cappelen Academic Publisher.

Martin, M. O., Mullis, I. V. S., Foy, P. (2008). *TIMSS 2007 International science report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mathisen, V. (1997). *Evaluering av "åpen bok" eksamen i grunnkurs økonomi og informasjonsbehandling, mekaniske fag og naturfag* (Evaluation of "open book" exams in Economy, Information management, Mechanical subjects and Science)ØFRreport no. 25. Lillehammer: Østlandsforskning.

Ministry of Education (1996). *The curriculum for the 10-year compulsory school in Norway*. Oslo: The Royal Ministry of Education, Research and Church Affairs.

Ministry of Education (1998). *Guidelines to the Curriculum for the 10-year compulsory school in Norway*. (L97). Oslo: The Royal Ministry of Education, Research and Church Affairs.

Ministry of Education (2006).*Curriculum for the Knowledge Promotion*. Provisional printed version. Oslo: The Royal Ministry of Education and Research.

Mullis, I. V. S., Martin, M. O., Foy, P. (2008). *TIMSS 2007 international mathematics report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

NOU (2003). *I første rekke. Forsterket kvalitet i grunnopplæring for alle (In first line. Enhanced quality in a basic education for all)* . Report No. 16. Oslo: The Royal Ministry of Education and Research.

OAGN (2006). *Riksrevisjonens undersøkelse av opplæringen i grunnskolen* (The Office of the Auditor

General of Norway's investigation of primary and secondary school), Document 3:10. Oslo: The Office of the Auditor General of Norway.

OECD (1988). *Reviews of National Policies for Education Norway*. Paris: OECD.

OECD (2010). *PISA 2009: Science Competencies for Tomorrow's World*. Paris: OECD.

Propper, C., Wilson, D. (2003). The use and usefulness of performance measures in the public sector. *Oxford Review of Economic Policy*, 19(2), 250-267.

Sejersted, F. (2005). *Sosialdemokratiets tidsalder* (The area of the social democracy). Oslo: Pax Forlag.

Skagen, K. (2001). *Pedagogikkens elendighet. Ukorrekte artikler på randen av litteratur. Essaysamling om skole og pedagogikk* (The tragedy of education. Essays on schooling and education). Kristiansand: Høgskoleforlaget

Skagen, K. (2004). Det nye pedokratiet i skoleverket (The new "educracy" in the school system), *Utdanning* April 2th, 2004

Skrede, S. (2004). Hvorfor er norske elever svake i fag og på topp i bråk? (Why are Norwegian students poor in school subjects and on top regarding noise and disorder?) *Aftenposten*, 20(12,) 11 (chronicle).

Telhaug, A. O. Mediås, O. A., Aasen, P. (2006). The Nordic Model in Education: Education as part of political system in the last 50 years. *Scandinavian Journal of Educational Research*, 50 (3), 245-283.

Tønnesen, L. K. B. (2004) *Norsk utdanningshistorie. En innføring med fokus på grunnskolens utvikling*, (Norwegian educational history). Høgskoleforlaget, Kristiansand.

UNESCO (2006). *International Standard Classification of Education*. Paris: UNESCO.

White Paper No. 37 (1990-91). *Mot rikare mål* (Towards richer goals). Report to Parliament. Oslo: The Royal Ministry of Education.

White paper No. 30 (2003-4). *Kultur for læring* (Culture for learning). Report to Parliament. Oslo: The Royal Ministry of Education and Research.

White Paper No. 31 (2007-8). *Kvalitet i utdanning* (Quality in education). Report to Parliament. Oslo: The Royal Ministry of Education and Research.

Wiggins, G. (1990). The case of authentic assessment. Downloaded from <http://eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?acno=ED328611>

W.W.C. (1922). Editorials. **Regulating the project**. *The Journal of Educational Research*, 5 (3), 245-246.

Advised by Kaare Skagen, Oslo University College, Norway

Received: *January 22, 2011*

Accepted: *March 17, 2011*

<i>Eyvind Elstad</i>	Professor, Department of Teacher Education and School Research, University of Oslo. Phone: + 0047 22855070. E-mail: eyvind.elstad@ils.uio.no Website: http://www.uv.uio.no/ils/personer/vit/eyvindel/index.html
<i>Are Turmo</i>	Assoc. Professor, Department of Teacher Education and School Research, University of Oslo. Phone: + 0047 22855070. E-mail: are.turmo@ils.uio.no Website: http://www.uv.uio.no/ils/personer/vit/aret/index.html
<i>Øystein Guttersrud</i>	Assoc. Professor, The Norwegian Centre for Science Education, University of Oslo. Phone: + 0047 22855050. E-mail: oystein.guttersrud@fys.uio.no Website: http://www.naturfagsenteret.no