Mandatory coursework assignments can be, and should be, eliminated!

John Haugan, Marius Lysebo and Per Lauvas

Faculty of Technology, Art and Design, Oslo and Akershus University College of Applied Sciences, 0130 Oslo, Norway

ARTICLE HISTORY
Compiled February 24, 2017

ABSTRACT
Formative assessment can serve as a catalyst for increased student effort and student learning. Yet, many engineering degree programs are dominated by summative assessment and make limited use of formative assessment. The present case study serves as an example on how formative assessment can be used strategically to increase student effort and improve student learning. Within five courses of an engineering bachelor degree program in Norway, the mandatory coursework assignments were removed and replaced by formative-only assessment. To facilitate the formative assessment, weekly student peer-assessment sessions were introduced. The main findings include an increase in student study hours and improved student performance on the examinations. Finally, interviews were conducted by an external consultant in an effort to identify key factors that attributed to the positive outcome.

KEYWORDS
formative assessment; peer-assessment; mandatory coursework assignments

1. Introduction

We have noticed that many of our colleagues spend substantial time and energy to control the students in an effort to 'force' learning. This is accomplished by excessive use of mandatory coursework assignments. A limited review of the course documentation at randomly selected Norwegian colleges' and universities' web sites, confirms that most of the courses in engineering degree programs have mandatory coursework assignments. The number of assignments varies in the range of 3 to 12 per semester per course. However, as many as 16 per course were found at our institution.

A fairly typical mandatory coursework assignment in mathematics, physics and related subjects, is made up of 2 to 10 problems from the textbook. The assignments are due in one or two weeks, they are reviewed by teachers or teaching assistants and are graded on a pass/fail scale. If a student fails one or several assignments, that student is not allowed to sit the final exam. Sometimes it is possible to resubmit a failed mandatory assignment, but at our institution it depends on the teacher. This somewhat 'messy' practice is only allowed because the assignments are judged on a pass/fail scale.

CONTACT J. Haugan. Email: John.Haugan@hioa.no

Clearly, the students need to do what it takes to pass these mandatory assignments. As a consequence, they have to work on a first thing first basis (Michelsen and Aamodt 2007). Some would argue that the mandatory coursework assignments help the students to avoid procrastination. However, when the number of assignments per course reaches two digit numbers, the cumulative 'forced workload' becomes exhausting. To meet a deadline, the students may need to postpone other study-related tasks. Subsequently they risk falling behind in several courses and develop 'holes' that can make it difficult to keep up with the current themes being lectured, especially when one or more assignments are disapproved and have to be resubmitted.

As the number of students grows, feedback becomes increasingly scarce and late. In fact, at our institution many students do not bother to pick up their assignments because the feedback, if any, comes too late to be useful. Stories of plagiarism and other undesirable strategies are not unheard of, although rarely reprimanded by faculty - probably due to the large number of assignments. The teachers (or teaching assistants) need to approve the assignments, but the resources needed to provide high quality feedback in due time to large classes are not available (Michelsen and Aamodt 2007). The result is an expensive and rather intractable feedback and assessment system for both students and faculty alike. The focus seems to be on doing rather than on learning, perhaps inspired by an idea that students who are working are also learning what they are supposed to learn.

It has to be true that it simply is impossible to learn without spending time on learning activities (Carroll 1963). It is also self-evident that not all learning activities are equally productive for effective learning to take place. The students have to allocate the time needed to learn what shall be mastered, and equally important, spend the time on task in appropriate ways that are conducive for effective learning.

Two key elements relate to feedback and to the way students are assessed (Black and Wiliam 1998; Hattie and Timperley 2007; Chickering and Gamson 1987; Gibbs 2010). To structure the discussion of assessment systems, it is common practice to distinguish between summative and formative assessment (Scriven 1967; Sadler 1989). The goal of summative assessment is to assess how much the students have learnt at the end of an instructional unit and to ensure that standards are met. In higher education, this is most often accomplished by marking students' work.

Formative assessment, on the other hand, is intended exclusively to serve as a purpose for learning. First, there are no marks associated with it. Second, rather than trying to cover up, hide or ignore topic areas that were not well understood, the students should bring problems and possible misconceptions front and centre to get as much help as possible. Whereas summative assessment needs to be accompanied by formal procedures, formative assessment is free from bureaucracy and can be provided by anyone knowledgeable.

Following the work of Graham Gibbs, assessment systems have six main functions (Gibbs 1999a, 23):

(1) capturing student attention and effort;
(2) generating appropriate learning activity;
(3) providing feedback to the student;
(4) developing within students the ability to monitor their own learning and standards;
(5) allocating marks - to distinguish between students or to distinguish degree classifications;
(6) ensuring accountability - to demonstrate to outsiders that standards are satisfactory.

All functions are important. The present work is, however, mainly concerned with items (1) - (4). These items are closely related to formative assessment, whereas the summative assessment system has to be responsible for items (5) and (6).

Mandatory coursework assignments capture student attention and effort. That is, they are very successful in supporting item (1) and may also support items (2), (3) and (4) depending on the setup. However, we claim that the extensive and uncoordinated use of assignments that we have described, does not. Seeing this, we argue that the current assessment system should be refurbished to better support learning. After all, items (2), (3) and (4) are just as important as item (1), if not even more so. Item (2) is somewhat ambiguous because different students benefit from different learning activities. Therefore it seems appropriate to let the students have some say in what they spend their time on (Ryan and Deci 2000; Felder and Brent 2005). Eventually they need to monitor their own learning and make choices accordingly, as emphasised by item (4). Needless to say, appropriate and useful feedback is essential in this respect.

It is our opinion that the two principles of prompt feedback and time on task are hardly met with the current practice. As we see it, the comprehensive use of mandatory coursework assignments

- makes it nearly impossible for teachers to give prompt feedback with the limited resources allocated to the different courses
- forces the students to work on a first-thing-first basis, making it hard for them to keep up with the themes currently being lectured.

These observations led us to ask the obvious question:

How can the courses be reorganised to ensure prompt feedback and increase student effort without teachers exceeding their allocated resources?

In short, two changes were made to five courses of an engineering degree program:

- All mandatory coursework assignments, except for the final exam, were removed.
- Weekly student peer-assessment sessions, based on weekly problem sets, were offered on a voluntary basis.

All together, 28 mandatory coursework assignments were removed from the courses. This may be viewed by some as a radical approach, and the questions we wanted to investigate are:

(1) Will student effort flatline throughout the semester and only peak around the time of the examinations?
(2) Will these changes impact the results of the examinations?

Failing rates in introductory courses in engineering education vary substantially at our Faculty, and any measures that can be taken to increase examination results would be highly appreciated.

2. Methodology of the study

The study originated in the context of teaching. We started out by giving increased attention to formative assessment. Only some minor elements in the courses were modified. The consequences were encouraging and the study was scaled up to cover 6 courses delivered to the students in the same program. Data was gathered about the entrance knowledge level in relevant subject matter, participation rate in the voluntary activities (i.e. formative assessment parts) and the final examination results. An independent, external resource person acted as a consultant in the teaching process of the two first courses, interviewing a sample of students in different phases, feeding back reports to the involved teachers about students responses, evaluations and suggestions. The latter was also a formative assessment exercise to be used to improve the teaching/learning process in the making. In the final summative assessment of the program, the learning trajectories of the students is focused with the qualitative data serving as information to gain understanding of the results.

The main purpose was to test the hypothesis that formative-only assessment applied in courses in the Bachelor program would give as good results as suggested by leading advocates in the field of contemporary assessment. Kerdijk et al. (2015) found the cumulative test encourages medical students to distribute their learning activities over a course. However, there was no evidence for any short-term effect on overall knowledge gain. van der Hulst, M. and Jansen, E. (2002) found that the spread of study activities over the year is an aspect of curriculum organisation that has effects on student progress. Jansen (2004) found that summative assessment distributed over the teaching period yielded better results that one final examination only. Our contribution will add with the results of a third alternative, namely formative-only assessment integrated into the teaching process with one final examination thereby spreading students learning activities over the teaching period on a voluntary basis.

3. Theoretical lens

It has repeatedly been shown that formative assessment has a strong and positive impact on student learning (Hattie and Timperley 2007; Black and Wiliam 1998; Page 1958; Crooks 1988). As the volume of formative assessment increases, students are more likely to adopt a deep approach to learning (Gibbs and Dunbar-Goddet 2007), i.e. try to make sense of the material rather than memorise as many details as possible. Enhanced motivation and development of metacognitive skills are also associated with formative assessment (White and Frederiksen 2000). Moreover, Black and William concluded that formative assessment helps the low achieving students in particular (Black and Wiliam 1998). This is still debated.

The relationship between 'intrinsic motivation', see e.g. the early work of deCharms (1968), and formative assessment appears to be particularly strong. Ryan and Deci (2000) describe intrinsic motivation as motivation to learn in 'the absence of external impetus'. Intrinsically motivated students learn because it is 'inherently interesting and enjoyable'. For educators it is important to be aware that the students' intrinsic motivation is not constant over time, neither is it the same for different subjects. It can be undermined or it may be maintained and increased. Student autonomy seems to be the key. Students faced with external control mechanisms such as close supervision, testing and evaluative pressure, will not maintain their intrinsic motivation over time (Niemiec and Ryan 2009).

Mandatory coursework assignments attempt to *combine* formative and summative assessment: It is summative because the students need to pass the assignments to sit the exam, and it is formative because comments to help the students learn and progress may be part of the assessment. From the available studies, as well as our own experience, we draw the conclusion that combined summative and formative assessment is widely used and rarely successful. The study of Butler (1988) cannot be read in other ways than as a signpost of the limited value of such combinations and a forceful recommendation to seek the alternative paths of formative-only assessment. We lean on studies that support this view, showing that comments only is more effective than comments combined with grades (Black and Wiliam 1998; Crooks 1988; Page 1958). From our point of view, the main problem associated with the current mainstream practice in Norwegian engineering education is the very combination of formative and summative assessment, even when the grading is limited to a dichotomous pass/fail. At the secondary education level, the general picture is the same (see Havnes et al. 2012, 23):

> The overall picture that emerges through the data is that teachers give feedback on tests and assignments in addition to grade. Feedback without grade ... or withholding grade till the students have attended to the feedback, is rare. Feedback is tied to grading.

We argue that the positive effects of formative-only assessment can be easily destroyed in the somewhat intricate balance between the certification purpose and the feedback purpose. For feedback to be permitted to have the desired benefits, it needs to be separated from the summative functions. Not only is it necessary to give priority to formative assessment since it is impossible to apply both (Gibbs and Dunbar-Goddet 2009). It is also mandatory to employ assessment that is formative only and distinguish it clearly from all summative assessment. Then there is a realistic hope to benefit fully from the potentially potent form of formative assessment.

Hattie and Timperley (2007) introduced four major levels of feedback: (1) feedback about a task; (2) feedback about a process; (3) feedback about self-regulation; (4) personal feedback. Feedback about a task is concrete and easily attainable. In subjects such as mathematics and physics, feedback about the process is primarily about solution strategies. According to Hattie and Timperley (see 2007, 93), self-regulation

> involves an interplay between commitment, control and confidence. It addresses the way students monitor, direct, and regulate actions toward the learning goals. It implies autonomy, self-control, self direction, and self-discipline.

Personal feedback, such as 'You are a great student', is nice but not effective to support learning (Kluger and DeNisi 1998).

The fostering of autonomy and self-regulation may be viewed as the ultimate objective of any feedback intervention, and 'good feedback practice is broadly defined here as anything that might strengthen the students capacity to self-regulate their own performance' (Nicol and Macfarlane-Dick 2006, 205).

Timely feedback about both the task and the process is found to be very powerful, in particular when (Earley et al. 1990; Hattie and Timperley 2007; Harackiewicz 1979; Harackiewicz, Manderlink, and Sansone 1984; Butler 1988):

(1) it is clear, purposeful and meaningful
(2) it corrects erroneous interpretations and ideas
(3) it supplies cues to develop new understanding beyond the task at hand

Note that successful feedback about tasks and processes can lead to increased self-regulation (Nicol and Macfarlane-Dick 2006). Increased self-regulation tends to increase student effort and engagement (Hattie and Timperley 2007).

Feedback timing has been studied extensively. Clariana, Wagner and Roher Murphy (2000) concluded that optimal feedback timing depends on the difficulty of the task. Delayed feedback is beneficial when the task is truly challenging and requires a sustained thought process. Feedback on a relatively simple task should be as immediate as possible, whereas it makes little sense to rush feedback aimed at promoting self-regulation.

4. Teaching model

Biggs (1996) and Biggs and Tang (2007) argue that courses in higher education should be based on the concept of constructive alignment. This concept is based on two principles:

(1) Constructive refers to the idea that everyone learns by constructing knew knowledge based on prior knowledge. It is what the learner does and thinks that matters, what the teacher does is second to this.
(2) Alignment means that there is a clear connection between 1) the description of the intended learning outcomes, 2) the learning activities the students take part in, 3) the teaching activities introduced by the teachers, 4) the exercises and the problem solving activities, and 5) the assessment methods, including how feedback is given. All five items have to fit together to maximise the learning outcome.

We have taken great care to build our teaching model on the principles of constructive alignment. This means for example that the weekly problem sets (see section 4.2) are constructed from the same intended learning outcomes as the problems given on the final exam, as are the content and the style of the lectures, exercise sessions and tutorials.

4.1. *Changes to the assessment system*

The assessment systems of five existing courses, all part of one engineering bachelors degree programme, were made formative-only by making the following simple changes:

(1) The mandatory coursework assignments were removed. These assignments used to be graded on a pass/fail scale, and the students needed to pass them to get access to the exams. However, the marks came *exclusively* from the final written examinations also before the intervention described in the present work.
(2) Weekly problem sets were introduced.
(3) Weekly peer-assessment sessions were introduced by allocating time from the periods previously designated to lecturing.

Key-information about the five courses is given in Table 1. Previously, all these courses were taught in a rather traditional way with lectures, exercise sessions and a number of mandatory coursework assignments.

Table 1.: Information about the courses described in the present work. The lectures and exercise sessions are 45 minutes long.

Courses	Semester	Lectures (per week)	Exercise sessions (per week)	ECTS credits	Number of students
Mathematics	first and second	2	2	5+5	42
Physics	first	2	2	5	46
Thermodynamics	second	4	2	10	49
Fluid Mechanics	third	4	2	10	65
Vector Calculus[a]	fifth	6	0	10	206

[a] Vector Calculus is an elective course offered to all students.

4.2. *Weekly problem sets*

The pillar in our effort to align the elements of the courses is the weekly problem sets. The clear alignment makes it possible to formulate this 'contract' articulated to the students:

> If you work seriously with the weekly problem sets, i.e. finish them and understand them, the probability that you have achieved this weeks learning outcome is great.

In the mathematics courses, weekly learning outcomes are formulated explicitly on the weekly problem sets. In the other courses, the learning outcomes are exemplified through the problems, but not stated explicitly. The difference is due to the teachers preferences.

We have taken great care to design the weekly problem sets so that they are aligned with the intended learning outcomes. This means that most of the problems could have been given on the final exam, making it worth the effort to understand them. A typical problem set contains 5 to 10 problems on different taxonomical levels. The problems are designed and written by the teachers. Some problems are traditional and require the usual quantitative calculations, others are qualitative (paradoxes, thought experiments, interpretation of diagrams etc.) and require reasoning and arguments. Occasionally it is necessary to write a few lines of computer code to solve the problems. The workload is adjusted to match the expected number of study hours in the courses.

The problem sets orient student effort towards the themes of the lectures. Understanding is put in the forefront to continuously remind the students that it is more important to formulate a problem mathematically and evaluate the results, than to focus on algorithmic and trivial procedural knowledge. This is important in itself (Ramsden 2003), but particularly important to make the peer-assessment sessions worthwhile.

4.3. *Peer assessment*

Approximately 20 minutes are allocated to the peer-assessment sessions every week. The students assess the work of a fellow student providing immediate feedback on it. The students are invited to take part on a voluntary basis with one limitation only. Only the students who have worked on the problems of the week are allowed to correct the work of other students. One of the reasons for doing this is to prepare the students for their professional lives, knowing that evaluation and assessment of their own and others' work are a central part of the engineering profession (Falchikov and Boud 1989). There are other potential benefits of peer assessment as well. Students seem to be competent to give formative feedback. More important, peer-assessment is associated with the development of critical thinking, increased responsibility for own progress, higher motivation and prompt feedback (Topping 1998; Ertl and Wright 2008; Boud and Falchikov 2007).

We have tried to moderately vary the form of peer-assessment as described below.

(1) Peer-assessment of solutions to the weekly problem sets

This particular form of peer assessment is very much based on Graham Gibbs presentation in (Gibbs 1999b). The students assess each others solutions to the weekly problem sets. The following set of rules were negotiated for these sessions:
- To get access to the peer assessment sessions, the students have to hand in their own solution during the 15 minutes break in the lecture.
- It is perfectly acceptable not to have succeeded in solving the problems.

However, it is a requirement to explain, in writing, why one or more of the problems were too hard to solve.

- The students' solutions are randomly distributed to fellow students at the start of the peer-assessment sessions.
- A selected problem from the weekly problem set is talked through by the teacher. During this revision, the students assess and comment each other's solutions to this one problem.
- Finally, the solutions are redistributed to the owner with corrections and comments from a fellow student. If the comments are disputed or deemed unsatisfactory, it is possible to get a second opinion from the teacher. This option was practically never used.

(2) Peer-assessment of in-class tests The students are given a new problem to solve in class. The problem is very similar to one or several of the problems on the current problem set, and it serves as a self test to check if the weekly learning outcomes are reached. The test problem is solved individually. After 5 10 minutes, the students hand over their solution to a fellow student sitting nearby. The teacher then solves and discusses the problem in a plenary while the students comment and correct their fellows solutions.

Finally, the students write a short assessment of both their own and their fellows performance on the test. The students are instructed to determine which parts of the solution that are satisfactory, and which parts that need improving. The aim is to make the students reflect over both their own and their fellows performance and learning. The assessments are handed in and read by the teacher (not assessed).

To get access to the peer assessment sessions of in-class tests there are no requirements. In the elective Vector Calculus course, the peer assessment sessions were based exclusively on short in-class tests. However, the students have to participate in a certain number of these peer-assessment sessions to sit the exam. The students enrolled in this course come from seven different bachelor degree programmes, and we are therefore not in a position to train the students in the same way as in the other courses.

Initially the model answers were made available to the students *only* after the peer-assessment sessions. However, the students did not appreciate this and the setup was changed. It also turned out to be hard for the students to do a thorough assessment of all the problems during the time they had at their disposal. Consequently we discontinued this practice and only ask them to assess one or two of the weekly problems. Also, not all problems are well suited for peer-assessment. We prefer problems that allow for different solution strategies, conceptual problems or problems that ask for written explanations. The quality of the comments that students write to each other vary. The teacher may attempt to improve the usefulness of the comments by reminding the students what to look for and to discuss with the students what high-quality work looks like. The students value these discussions and the opportunity to see how fellow students solved the problems highly. Students take part in the peer-assessment scheme for these two reasons, the comments they receive are regarded as less important.

4.4. *Summative assessment*

The part of the summative assessment system responsible for the grades is kept unchanged. The students take one 3 hours individual written exam at the end of the semester, i.e. in June or December. The grades are determined exclusively on the results of this exam. All printed and hand-written material, as well as a calculator, may be used. The ECTS grading scale is applied with six levels where 'A' is the best mark, indicating a superior understanding of the subject, and 'F' is fail.

5. Results

The present case study includes 320 *unique* undergraduate engineering students. The participants took one or several of the regular examinations in the five courses described in Table 1.

5.1. *Time and effort*

We asked the students to estimate the total number of hours that they spend on average per week per course. The results of these surveys are given in Table 2. Note that the surveys are done in class, with a fairly high response rate (65 - 90 %), but with the possibility to be positively skewed. Setting aside the question of whether or not study-hours data are reliable, we argue that the numbers in Table 2 nevertheless are notable since they were collected in the same way as during earlier years.

Table 2.: The total number of hours spent on academic work on average per week as reported by the students.

Course	Total number of study hours	ECTS credits	Norm (hours)
Mathematics	6.60 (+ 2.52)	5+5	6.67
Vector Calculus[a]	14.0 (+ 2.30)	10	13.3
Physics	6.96 (+ 0.91)	5	6.67
Fluid Mechanics	13.5 (NA)	10	13.3

[a] Numbers in parentheses indicate differences from the previous year. Data from previous years for Fluid Mechanics is missing. The 'norm' has been established on the assumption that 30 ECTS credits correspond to a standard 40 hours working week.

First, compared to the results of similar surveys conducted in the past, the students now report that they study more. In particular we note a substantial increase in student effort in both the Mathematics and Vector calculus courses. A recent survey (2016) indicates that the number of study hours in the Mathematics course keeps increasing. (8.0 hours per week on average, although dropping to 7.2 hours two months later.)

Second, the Norwegian Agency for Quality Assurance in Education (NOKUT) conducts the annual national student survey for higher education in Norway. Their data show that the average student, on this particular degree programme, spent 33.4 hours per week on academic work (2015), including voluntary cooperation with other students (NOKUT n.d.). 75% of the students participated in this particular online survey. It means that the weekly student effort, on average, is well below the norm (40 hours per week). We have, on two separate occasions, asked the same question in our surveys, and the results are similar (35.9 hours and 31.2 hours). In light of this, the data presented in Table 2 suggest that the weekly study hours are comparatively high within these courses.

The reader may well ask why the students would choose to work more as a response to the changes made? We acknowledge that the answer to this question is not simple and straightforward. However, the students themselves provided some clues:

> You teachers play around with some psychology here: Because we don't have to do the problem sets, it's more fun to do them.
>
> I don't have to spend time on problems that I know how to solve.
>
> I work hard because I am afraid to fall behind.

The first quote is the one most often encountered in our student surveys.

Some of the games played by some students in the standard course layout with frequent obligatory assignments neither seem to be necessary nor tempting to resort to. 'Faking good' (Gibbs 2006) or copying someone's work, have no place in a formative-only assessment system. One could therefore speculate that not only do the students work more, but they also spend their time on more worthwhile activities than before. However, we do not have hard data to support this view.

5.2. *Exam results*

With no mandatory assignments, the students only need to turn up on the day of the exam. No attempts are made to control that the students are qualified before they sit the exam. The examination results for the courses are presented in Fig. 1. Students who fail are entitled the right to resit the exam the following semester. The data presented in Fig. 1 are limited to the results from the ordinary examinations.

It is common experience that the introductory course in mathematics can be a daunting experience for many engineering students. Failure rates exceeding 50% are not unheard of. In addition, the fraction of students that pass with an above average mark can be very low. The data presented in Fig. 1 tell a different story: Only 14.3% of the students failed and 16.3% were awarded 'A'. In addition, more than half the class were awarded a 'C' or better. The year before the intervention, 42% failed and 33% received an 'E' or a 'D'.

From 2012 to 2014, 25% of the students failed the fluid mechanics course and 35% received an 'E' or a 'D' (three year average). Two ordinary examinations have been arranged in that course since the aforementioned changes were implemented, and well below 10% have failed. Figure 1 includes the two-year average grade distribution for the sake of comparison. However, the reader is warned that the comparison may be unjust because, in addition to the changes described in Sec. 4, the lecturer was also replaced.

Although it is impossible to prove with rigor, we presume that the students perform better because they have learned more. The exams have all been reviewed and approved by internal or external examiners, and we believe that the quality of the exams is sufficiently high to reliably measure the students knowledge. However, a new exam is made every year, and the impacts of the inescapable variations are unknown. In addition, a new group of students, with slightly different abilities, interests and informal prerequisites, is enrolled every year. No systematic effort has been made to investigate the effects of any of these variables.

The results in the courses we have not commented explicitely are also good, but the improvements in grades are modest and within the 'normal variations'. Albeit the trend is positive, more data is needed to attempt a conclusion.

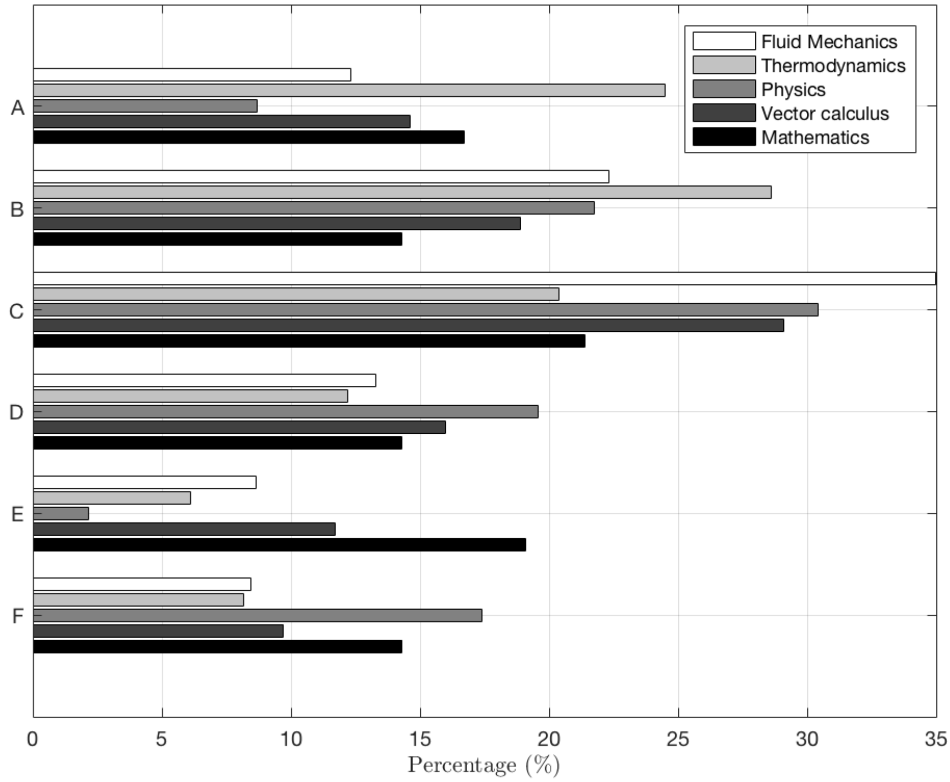

Figure 1.: Distribution of grades from the ordinary exams in the five courses Mathematics, Vector calculus, Physics, Thermodynamics and Fluid Mechanics. The distribution of grades in the Fluid Mechanics course is the average over two years. See the text for more details.

5.3. *Evaluation by an external consultant*

An external consultant was employed to evaluate the mathematics and physics courses. The external consultant interviewed 15 students and followed the lectures regularly. The external consultant also sat in on the exercise sessions occasionally. All interviewees were guaranteed anonymity and the interviews were voice recorded and transcribed.

Early in the first semester, the interviews focused on how the informants engaged with the course material with special reference to peer-assessment. A few weeks into the semester, informants were selected from a group of students that chose not to participate in the peer-assessment sessions. Towards the end of the second semester, a third round of interviews were conducted. In these interviews the informants were asked to evaluate the two different modes of peer-assessment that they had experienced. (Both forms are described in Sec. 4.3.)

The teaching model was very well received by the students. The teachers enjoyed high credibility among all the informants, and the weekly problem sets were regarded as very important by most, if not all. There seemed to be wide agreement that the goals and standards were clearly communicated. It stands to reason that this positive attitude towards the courses made the students more inclined to study hard:

> I put a lot of hours in to understand the material, because I need to. It is very important for me to solve the exercises. My goal is to figure out how to solve them, even when it means banging the head against the wall for hours. I go to all the lectures and all the exercise sessions. Still, it is important for me to get access to the model answers. I prefer to work alone and independently. Working together with my peers is not for me.

No two students are alike, and peer assessment may not fit the particular needs of all. The students should be secured the right to make their own choices as to what suits them best within the space available to them. We have found it to be reasonable to require students to come prepared for the peer assessment sessions. Some students find it unreasonable to be excluded in this way. The students' reasons for endorsing the merits of peer assessment vary accordingly as indicated in some student responses:

> It is essential to get feedback on your own work as quickly as possible. It is also important to give feedback; one learns much from giving feedback to others and to see what they have done, having to judge what is right and what is wrong.
>
> To me it is both fun and productive to work alone as well as together with fellow students. It is a good thing to get training in feedback, both in the role of the receiver and provider. The workload put on the students in this place, is quite demanding and the students who do not manage to keep up with the pace, will encounter problems. Probably, the attrition will be highest during the first semester and one has to learn to prioritise from the very start.
>
> Spending time on working on the problems, assessing one's work and the work of your fellow students must be given high priority which is the case here if you only take part in what is provided for you.

Even though the students are very content with the courses, they rank the teaching activities differently, in particular the peer-assessment sessions. The majority of students value the opportunity to see how other students solved the problems. However, the quality of the feedback from their peers was regularly questioned.

Initially, practically all the students chose to participate in the peer-assessment sessions. At that time, these sessions were only about the problem sets, i.e. form 1 as described in Sec. 4.3. After 8 weeks, 10 – 20 % of the students regularly chose to leave the classroom when the peer-assessment session began. The external consultant interviewed some of these students. The interviews suggest that they can be divided into two groups: Those who did not need to attend, and those who did not have the time to attend. In the first group we find students that attended whenever needed, i.e. when the material was not well understood. These students were very selective with what they spent their time on. This behaviour was encouraged by faculty. In the other group, we find students that could not allocate enough time to complete the weekly problem sets. This was sometimes due to external factors such as their family situation or because they held a part time job outside of academia. On other occasions, it was due to project work or mandatory assignments in courses that run parallel with the courses described in the present work.

Two different modes of peer assessment were described in Sec. 4.3. The majority of our students found peer-assessment of in-class tests more effective than peer-assessment of the weekly problem sets. At their best, the tests function as superb indicators of whether or not the core material was well understood. Furthermore, the students could get help and receive feedback on their solutions to the weekly problem sets in the exercise sessions, thereby making the peer-assessment scheme appear somewhat redundant. As emphasised and greatly appreciated by most informants, model answers posted online were made available. The social pressure that accompanies both modes of peer-assessment was regarded as acceptable by the informants although not highly appreciated by all, in particular those with tight time budgets.

In some of the courses, data from a voluntary entrance test were compared with the examination results and the participation rate in the formative assessment. When looking at individual trajectories, the general trend is obvious. Students with excellent entrance knowledge of the subject and who took part in formative assessment, ended up with the best examination results. However, students starting with low score on the entrance test could possibly, at least in some cases, perform far better than could be expected. Formative only assessment seemed to be most profitable for them. Given that entrance qualifications vary substantially among the first-year students, excellent results by the students with low entrance scores, provides a solid platform for further studies.

Some examples: In the physics course, half the students with low scores (below 30%) on the initial test failed. However, one student started with a low score (13%) and obtained a B, 1 started with 17% and ended up with a C. In the mathematics course, almost half of the students who obtained an A, started with low scores on the initial test.

From the analysis the picture emerges that half of the students in lowest score bracket (≤ 19%), also obtain a low mark, but among the others, there are some students with extraordinary progress. One of them explained:

> In my secondary school in a small town, the quality of science teaching was not so high and my results not so good. My background when starting here, could definitely have been better. My progress has, however, been remarkable because the teaching is so good. I was not at all sure about my prospect but gained a positive assurance quickly.
>
> It is really enjoyable to perform well! You only go along and follow the instructions. Even when it is really difficult you get all the assistance you need to succeed. I enjoy the thoroughness. It is considered more important to gain in understanding than getting through as much stuff as possible. That is really good.

6. Summary and concluding remarks

The present work challenges the well-worn claim that mandatory coursework assignments that are to be approved or graded are required in order to make students work at a steady pace and to perform well. In a modularised system, the opposite seems to be the case. Mandatory coursework assignments tend to force the students to spread their effort unevenly across weeks, topics and courses. The weekly student effort fluctuates because the assessment system inadvertently prevents the students from working at a steady pace in all the courses. Being forced to work on a first-thing-first base, the students are left with few options to design their own work schedule.

The courses described in the present work have been moderately redesigned to induce some degree of autonomy and sound study approaches to first-year students and to contribute to good results. The ultimate goal is to make the students self-regulated learners who can monitor their own progress. The forms of delivery was kept unaltered, only some minor modifications in the assessment process were put in place.

Feedback is of the essence in any training program. A successful format is to have students work reasonably hard, evenly distributed from the beginning to the and, applying effective learning strategies in their work. Equally important is to have a sound assessment system that both is helpful in the learning process as well as providing adequate quality control of the students' learning results. Feedback is often linked to summative assessment, often in the form of coursework assignments. The results can be that the potential benefits of feedback are not utilized resulting in poor results, especially among the students who start with a weak knowledge base and a shaky self confidence.

We have tried to look into the potential value of feedback that is not linked to summative assessment, but as a learning propeller in its own right with the sole purpose of promoting and facilitating learning. Peer-assessment is certainly not the only alternative but had advantages, both as to potentials and practical ease, not the least when teachers' workload is concerned. With reasonably short sessions and by letting the students compare their own solutions to the weekly problem sets with the model solutions, the system could fairly easy be accommodated in the regular teaching program. The feedback had to be formative *only* and the students were given some access to decide upon the best practical arrangement and to decide for themselves which assignments to complete and how and when they wanted feedback. Our findings show that:

(1) the average number of study hours have increased, compared to previous years, in all the courses.
(2) the students performed very well on the examinations.

Initially our hypothesis was that these findings could be explained by the students' participation in the peer-assessment sessions. However, our students told a different story. They described an intrinsic motivation that they attributed primarily to the elimination of the mandatory coursework assignments, not to the peer-assessment sessions or anything else that was going on. Turning our initial hypothesis down, we now strongly believe that the clear distinction between formative and summative assessment made it possible for the students to focus exclusively on learning the subject matter. Misunderstandings and misconceptions were confronted in a supportive learning environment without any risk of sanctions from faculty. 'Faking good' or copying other's answers to exercises turned out to have no place in the assessment model.

Clearly, the present case study presents results for a situated case. Only two teachers have incorporated the main ideas in their courses. Notably, a course consists of many components, e.g. lectures, peer assessment sessions, exercise sessions, tests, an exam etc, and we have not attempted to isolate the impact of each component. Rather we have to conclude that the sum of the components support student learning as a whole.

References

Biggs, John. 1996. "Enhancing teaching through constructive alignment." *Higher Education* 32 (3): 347–364. http://dx.doi.org/10.1007/BF00138871.

Biggs, John, and Catherine Tang. 2007. *Teaching for Quality Learning at University (Society for Research Into Higher Education).* 3rd ed. Berkshire: Open University Press.

Black, Paul, and Dylan Wiliam. 1998. "Assessment and Classroom Learning." *Assessment in Education: Principles, Policy & Practice* 5 (1): 7–74. http://dx.doi.org/10.1080/0969595980050102.

Boud, David, and Nancy Falchikov. 2007. *Rethinking assessment in higher education: Learning for the longer term.* London: Routledge.

Butler, Ruth. 1988. "Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance." *British Journal of Educational Psychology* 58 (1): 1–14. http://dx.doi.org/10.1111/j.2044-8279.1988.tb00874.x.

Carroll, John B. 1963. “A model of school learning.” *Teachers College Record* 64 (8): 723–733.

Chickering, Arthur W, and Zelda F Gamson. 1987. “Seven principles for good practice in undergraduate education.” *AAHE bulletin* 3: 7.

Clariana, Roy B., Daren Wagner, and Lucia C. Roher Murphy. 2000. “Applying a connectionist description of feedback timing.” *Educational Technology Research and Development* 48 (3): 5–22. http://dx.doi.org/10.1007/BF02319855.

Crooks, Terence J. 1988. “The Impact of Classroom Evaluation Practices on Students.” *Review of Educational Research* 58 (4): 438–481. http://dx.doi.org/10.3102/00346543058004438.

DeCharms, Richard. 1968. *Personal causation: the internal affective determinants of behavior.* New York: Academic Press.

Earley, P. Christopher, Gregory B. Northcraft, Cynthia Lee, and Terri R. Lituchy. 1990. “Impact of Process and Outcome Feedback on the Relation of Goal Setting to Task Performance.” *The Academy of Management Journal* 33 (1): 87–105. http://www.jstor.org/stable/256353.

Ertl, Hubert, and Susannah Wright. 2008. “Reviewing the literature on the student learning experience in higher education.” *London Review of Education* 6 (3): 195–210. http://www.ingentaconnect.com/content/ioep/clre/2008/00000006/00000003/art00002.

Falchikov, Nancy, and David Boud. 1989. “Student Self-Assessment in Higher Education: A Meta-Analysis.” *Review of Educational Research* 59 (4): 395–430. http://rer.sagepub.com/content/59/4/395.abstract.

Felder, Richard M., and Rebecca Brent. 2005. “Understanding Student Differences.” *Journal of Engineering Education* 94 (1): 57–72. http://dx.doi.org/10.1002/j.2168-9830.2005.tb00829.x.

Gibbs, Graham. 1999a. “Improving Teaching, Learning and Assessment.” *Journal of Geography in Higher Education* 23 (2): 147–155. http://dx.doi.org/10.1080/03098269985425.

Gibbs, Graham. 1999b. “Using assessment strategically to change the way students learn.” In *Assessment matters in higher education: choosing and using diverse approaches*, edited by Sally Brown and Angela Glasner, Chap. 4. Buckingham: The Society for Research into Higher Education & Open University Press.

Gibbs, Graham. 2006. “How assessment frames student learning.” In *Innovative assessment in higher education*, edited by Cordelia Bryan and Karen Clegg, Chap. 2. London: Routledge.

Gibbs, Graham. 2010. *Dimensions of quality.* York, UK: Higher Education Academy. http://www.heacademy.ac.uk/resources/detail/evidence_informed_practice/Dimensions_of_Quality.

Gibbs, Graham, and Harriet Dunbar-Goddet. 2007. *The effects of programme assessment environments on student learning.* Oxford: Higher Education Academy. https://www.heacademy.ac.uk/sites/default/files/gibbs_0506.pdf.

Gibbs, Graham, and Harriet Dunbar-Goddet. 2009. “Characterising programmelevel assessment environments that support learning.” *Assessment & Evaluation in Higher Education* 34 (4): 481–489. http://dx.doi.org/10.1080/02602930802071114.

Harackiewicz, Judith M. 1979. “The effects of reward contingency and performance feedback on intrinsic motivation.” *Journal of personality and social psychology* 37 (8): 1352.

Harackiewicz, Judith M, George Manderlink, and Carol Sansone. 1984. “Rewarding pinball wizardry: Effects of evaluation and cue value on intrinsic interest.” *Journal of Personality and Social Psychology* 47 (2): 287.

Hattie, John, and Helen Timperley. 2007. “The Power of Feedback.” *Review of Educational Research* 77 (1): 81–112. http://rer.sagepub.com/content/77/1/81.abstract.

Havnes, Anton, Kari Smith, Olga Dysthe, and Kristine Ludvigsen. 2012. “Formative assessment and feedback: Making learning visible.” *Studies in Educational Evaluation* 38 (1): 21 – 27. http://www.sciencedirect.com/science/article/pii/S0191491X12000193.

Jansen, Ellen P. W. A. 2004. “The influence of the curriculum organization on study progress in higher education.” *Higher Education* 47 (4): 411–435. http://dx.doi.org/10.1023/B:HIGH.0000020868.39084.21.

Kerdijk, Wouter, Janke Cohen-Schotanus, B Florentine Mulder, Friso L H Muntinghe, and Ren A Tio. 2015. "Cumulative versus end-of-course assessment: effects on self-study time and test performance." *Medical Education* 49 (7): 709–716. http://dx.doi.org/10.1111/medu.12756.

Kluger, Avraham N., and Angelo DeNisi. 1998. "Feedback Interventions." *Current Directions in Psychological Science* 7 (3): 67–72. http://dx.doi.org/10.1111/1467-8721.ep10772989.

Michelsen, Svein, and Per Olaf Aamodt. 2007. *Evaluering av kvalitetsreformen.* Oslo: Norsk forskningsråd. http://www.forskningsradet.no/no/Publikasjon/Evaluering_av_kvalitetsreformen/1166550892135.

Nicol, David J., and Debra Macfarlane-Dick. 2006. "Formative assessment and self-regulated learning: a model and seven principles of good feedback practice." *Studies in Higher Education* 31 (2): 199–218. http://dx.doi.org/10.1080/03075070600572090.

Niemiec, Christopher P., and Richard M. Ryan. 2009. "Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice." *Theory and Research in Education* 7 (2): 133–144. http://tre.sagepub.com/content/7/2/133.abstract.

NOKUT. n.d. "Studiebarometeret." http://www.studiebarometeret.no/en/Sammenligne/257_HINGENERGI. Accessed: 2016-03-20.

Page, Ellis Batten. 1958. "Teacher comments and student performance: A seventy-four classroom experiment in school motivation." *Journal of Educational Psychology* 49 (4): 173–181. http://dx.doi.org/10.1037/h0041940.

Ramsden, Paul. 2003. *Learning to teach in higher education.* London and New York: RoutledgeFalmer.

Ryan, Richard M., and Edward L. Deci. 2000. "Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions." *Contemporary Educational Psychology* 25 (1): 54 – 67. http://www.sciencedirect.com/science/article/pii/S0361476X99910202.

Sadler, D. Royce. 1989. "Formative assessment and the design of instructional systems." *Instructional Science* 18 (2): 119–144. http://dx.doi.org/10.1007/BF00117714.

Scriven, Michael. 1967. *The methodology of evaluation*, 39–83. AERA Monograph series on evaluation. Chicago: Rand McNally.

Topping, Keith. 1998. "Peer Assessment Between Students in Colleges and Universities." *Review of Educational Research* 68 (3): 249–276. http://rer.sagepub.com/content/68/3/249.abstract.

van der Hulst, Monique, and Ellen Jansen. 2002. "Effects of curriculum organisation on study progress in engineering studies." *Higher Education* 43 (4): 489–506. http://dx.doi.org/10.1023/A:1015207706917.

White, Barbara Y., and John R. Frederiksen. 2000. "Metacognitive facilitation: An approach to making scientific inquiry accessible to all." In *Inquiring into Inquiry Learning and Teaching in Science*, edited by Jim Minstrell and Emily H. van Zee, 331–370. Washington, DC: American Association for the Advancement of Science.