

From Teaching as Transmission to Constructive Alignment: A Case Study of Learning Design

Vidar Gynnild, Bernt Johan Leira, Lars Erik Holmedal, Dag Myrhaug, Jon Coll Mossige,
Norwegian University of Science and Technology (NTNU)

ABSTRACT: This paper reports on an action research study on curriculum design in an optional module at the Norwegian University of Science and Technology. There were two professors teaching their respective parts, henceforth termed “Part 1” and “Part 2”. Oral evaluations confirmed students’ satisfaction with teaching; however, failure rates had been an issue for years and was a concern to the professors who were wondering why. This study sets out to explore causes of poor academic achievement in order to launch targeted interventions. A major issue was students’ limited capabilities in dealing with conceptual and theoretical problems. This was partially explained by a mismatch between types of problems in the exercises compared with those at the final exam. The ensuing intervention included the redesign of tasks to improve alignment of course components as well as introducing a mid-term exam specifically addressing conceptual and theoretical themes. Failure rates dropped markedly in subsequent years due to students’ enhanced ability to cope with the broader spectrum of course requirements.

Keywords: Course design, assessment, constructive alignment, science, engineering

1 AIM AND PROBLEM

Universities are under constant pressures to raise the quality of their offerings with the ultimate purpose of improving student learning to meet educational requirements and to equip candidates with competencies for their future careers. Such demands are dealt with by academics as well as by institutional quality assurance systems aspiring to improve educational quality by the application of data gathering tools to identify areas in need of attention. Consequently, the institution under investigation also had its own quality assurance system in operation, and one would expect that poor academic performance would be addressed in due time. The application of student reference groups would seem as a fitting arena to address shortcomings and introduce measures as required. However, since the predominant focus of those groups was on teaching rather than on learning, critical issues in need of attention escaped from the agenda. Rather, positive feedback from students served as re-assuring statements to the satisfaction of quality managers. Unfortunately, responses concealed academic shortcomings and failed to address critical aspects of the course.

The purpose of this study is first to explore learning outcomes by analyzing performance scores as seen at the exam; and second, to explain potential achievement patterns by examining the course structure: Why did students succeed academically in Part 1, while not to the same extent in Part 2? Based on insights, what might be fitting interventions, and to what extent did they make a difference to learning outcomes? The study aims to identify components impacting learning by practicing constructive alignment (Biggs, 1996). While “teaching” often is understood in a narrow sense as lecturing course content, the authors adopted a broader conception of the term including any measure contributing to learning (Goodyear, 2015). Theorizing connections between the learning environment and students’ academic achievement were subsequently used as a steppingstone to design measures to improve learning (Simmons & Gregory, 2003).

2 CONTEXTUAL SETTING

The semester structure featured four parallel courses, each with an estimated workload of 7.5 credits. The course under scrutiny is termed Stochastic Theory of Sealoads (henceforth “Sealoads”) and was taught in the fourth year of a five-year MSc study program. Essential course themes to be addressed

were Monte Carlo simulation, probability distributions for response, parameter-estimation, extreme-value statistics, stochastic processes, auto- and cross-correlation functions, spectra and cross-spectra, differentiation of stochastic processes, excitation-response of physical stochastic processes and response-statistics. Students should have detailed knowledge of principles and methods used to describe stochastic processes, including simple calculations of stochastic sealoads. They should also master the concepts and terminology used in statistical methods and in the description of stochastic processes.

Teaching methods were lectures, tutorials and problem-solving exercises, a model widely used in science and engineering education. The course served a dual purpose to prepare for engineering practice as well as to promote the acquisition of theory related to physical phenomena. This was an ambitious combination since new concepts and themes were introduced every week. Furthermore, due to speedy progression and heavy workload, strict priorities had to be made between academic and leisure activities.

Student enrolment varied by year, typically in the range of 30-70 annually. Course content carried great relevance in terms of employability, so recruiting students was never an issue. A particularly challenging feature was the diverse nature of its two parts, independently taught by the respective professors, who possessed long-term academic experience. The first part focused on statistical methods, while the other addressed stochastic processes, henceforth referred to as Part 1 and Part 2. While the first part largely re-addressed and extended key concepts and themes from a first-year statistics course, Part 2 introduced new concepts and themes of which students possessed limited or no prior knowledge. This challenged their ability to engage in deep approaches to learning (Marton & Säljö, 1976a, 1976b), and repeated patterns of poor academic performance in Part 2 motivated further exploration of potential causes.

3 THEORETICAL BACKGROUND

This study draws on a fundamental distinction between teaching and learning (Barr & Tagg, 1995). In the cited article, the authors termed the traditional, dominant approach to teaching the “Instruction Paradigm”. They concluded that too often means are confused with ends, and that learning rather than teaching should be the targeted goal in education. Their criticism was rooted in the fact that conceptions of teaching in the old paradigm were ineffective with students attending as passive listeners rather than active learners. The Learning Paradigm, as it was named, envisioned the institution itself as a learner over time, continuously attempting to draw on experiences to rethink and re-adjust its own practices to the benefit of students. To many of us this may seem a highly self-evident position; however, not necessarily so for stakeholders in all of higher education. Historically, many well intended measures have failed to make a difference to learning because they were rooted in conceptualizations of teaching as an end rather than a means. Criteria for success were related to student satisfaction, while the relationship between teaching and learning remained unaddressed.

The Learning Paradigm is an example of a Threshold Concept, enabling users to “make sense” and integrate many previously unrelated ideas within a certain domain. Such concepts serve to enhance individuals’ communication within disciplinary areas, and to extend practices in accordance with unique ways of seeing and experiencing the world around them. It may be compared with the figure-ground problem in psychology. Once aware of the figure, you never return to focus on the ground. The Learning Paradigm comes with characteristics distinctly different from the Teaching Paradigm. Terms used are “transformative” and “irreversible” learning (Land, Cousin, Meyer, & Davies, 2005). Such concepts come with an understanding of phenomena within a framework of thinking facilitating the integration of new ideas under an umbrella of principled reasoning, often termed a “gateway” or “portal”.

In a seminal article, the author explored the argument that “teaching in higher education will necessarily shift the balance of its efforts towards a greater investment in design” (Goodyear, 2015, p. 27). Ways of building design capacity is discussed with a focus on students’ activity with the ultimate purpose of helping them to take control over their own learning activities. With obvious implications for the current study, distinctions are made between three interpretations of teaching; interactive forms, pre-active forms and post-active forms, each featuring their unique purposes. While this cannot be the place for in-depth deliberations of each of them, the study draws on teaching as design as a relevant approach, emphasizing the alignment of course components to promote and ensure intended learning outcomes.

John Biggs combined ideas of instructional design with constructivist learning theory (Biggs, 1996), commonly known as constructive alignment, a framework frequently used to guide decision-making at all stages of course design. It draws on constructivist learning theory with an emphasis on alignment of course content, teaching and learning activities and assessment. Course design typically draws on theories of teaching and learning; however, such theories are often implicit. Theoretical underpinnings are taken for granted, while improvements often end up as fine tuning of existing practices. The authors acted on the belief that learning activities needed to actively promote intended learning outcomes.

Science and engineering courses can be characterized by speedy progression and heavy workloads, making course planning challenging and sometimes rather unpredictable. A serious attempt to understand the complexities of higher education environments was made in a seminal book published 50 years ago (Snyder, 1971). The author examined how students adapted to learning environments, with a view to inform change strategies: “As I examined ... I was struck repeatedly with the importance of a hidden agenda, a hidden curriculum, which determined to a great degree the way in which the various participants played the game, read the cues, adapted to their immediate educational circumstances” (Snyder, 1971, p. xii). Educational planning under such conditions is complicated because students react differently under identical environments.

If we think of learning as a change in the ways in which something is seen or experienced, much effort in higher education fails to achieve this goal. As Snyder correctly noted, some students translate learning tasks into set of procedures which “... may or may not have much to do with learning, or even knowledge” (Snyder, 1971, p. 6). This dissonance between the formal curriculum and what is required to pass the exam “... infer as the basis for the rewards in the particular setting” (Snyder, 1971, p. 6). Educational planning cannot be left at an intentional level but needs to be monitored and theorized to understand how variables combined aid to produce intended as well as unintended learning outcomes.

4 METHOD

The first author had served years of practice as an educator before joining the current study and was acquainted with the professors in charge of the course. The immediate motivation for a closer, longer term collaboration was rooted in concerns related to discontent with students’ academic performance, particularly as seen in Part 2 of the course. Jointly the authors formed a project with the stated purpose of improving outcomes by exploring student behaviors in relation to course design decisions. This combination of research and action is commonly known as action research and is widely used, and the literature offers rich accounts of the method’s suitability for learning and change (Zuber-Skerritt, 2002). An apparent benefit is the combination of research evidence and interventions. Actions are not random but draw on understandings of how different variables are connected.

Action research involves learning from experiences, and from critical reflection on experiences as a foundation for action. While drawing on Grounded Theory (GT), Grounded Action (GA) extends beyond the construction of theory by suggesting actions targeted to solve problems in given contexts. GA was originally used to address complex institutional and social problems but has later proven beneficial in other settings. The authors started by analyzing scoring data at exams in Part 1 and Part 2, noticing that annual scores featured characteristic patterns over the years, which in turn became a matter of exploration. Data collection started in 2010 and went on all through 2017 with continuous and informal processes of theorizing connections as annual scoring patterns called for explanations. Interventions were grounded in data and assigned exam scores were analyzed to examine whether actions served to achieve their intended purposes.

5 INTERVENTIONS AND FINDINGS

When examining the scoring and grading policy, it appeared that a pass (minimum 41 of 100) could be obtained by specializing in Part 1 at the expense of Part 2, or vice versa. Furthermore, problems posed in Part 2 were of a different kind compared with those in Part 1. This might at times cause insufficient alignment of problems in exercises and at the exam in Part 2. While exercises to a larger extent required algorithmic skills, exam items addressed theoretical and conceptual problems. Conversely, in Part 1 tutorials and exams addressed similar types of algorithmic problems.

The project featured three major interventions, particularly to enhance conceptual understanding in Part 2. First, the assessment regime was changed in 2013 making it mandatory to achieve a pass in both Part 1 and Part 2. Candidates could no longer practice instrumental strategies to sub-optimize achievements by addressing only one part. In 2016, a voluntary mid-term exam was introduced, addressing exclusively theoretical and conceptual themes. Achieved scores accounted for 30 % towards the grade, while those who skipped the mid-term could achieve a maximum of 70% by attending the end-of-term exam. Furthermore, readings and exercises were revised to meet requirements of constructive alignment. In effect, this implied a more even balance in both parts between conceptual and algorithmic problems.

As seen in the preceding paragraph, the project dealt not only with improving teaching in a narrow sense, for example by refining the professors’ presentations skills. Rather, the project team adopted a broader perspective, incorporating all vital parts of the curriculum. Crucially, it was observed that assessment drives learning (Rust, 2002). Assessment formats and types of problems operate in tandem to direct and reinforce learning processes, a fact frequently overlooked in course revision. In folk rhetoric the quest is typically for innovative assessment formats, while deliberations on the nature of tasks and assignment is often omitted. In this project, the role of problem design and constructive alignment appeared as urgent. In Part 2 previous exam items were added to the exercises to familiarize students with this type of problems. Table 1 summarizes results of the redesign process to meet requirements of constructive alignment to support a broader spectrum of targeted learning outcomes.

Table 1. Current and revised problem design in Sealoads

	Part 1	Part 2
Current Exercise Design	Emphasis on calculations	Emphasis on calculations
Current Exam Design	Emphasis on methods & calculations. Tasks closely linked to exercises	Emphasis on concepts and theory. Tasks not necessarily linked to exercises
Revised Exercise Design	Emphasis on concepts and calculations as seen in Intended Learning Outcomes (ILOs)	Emphasis on concepts and calculations as seen in Intended Learning Outcomes (ILOs)
Revised Exam Design	Tasks aligned with Intended Learning Outcomes (ILOs)	Tasks aligned with Intended Learning Outcomes (ILOs)

Figure 1 portrays number of students by year (bottom), percentage of students failing the course and mean grades (top). Assigned scores ranging from 0-100 were translated into letter grades according to agreed conversion rules A (89-100); B (77-88); C (65-76); D (53-64); E (41-52); F (0-40). In Figure 1, the hovering boxes at the top visualize score boundaries for the respective grades. Assigned scores result in the mean grade “C” for all except two years, 2013 and 2014. However, failure rates vary dramatically over the years. Prior to introducing the pass-both-parts rule in 2013 failure rates were kept in the range of 5-20% annually. The year 2013 marks a change from the application of an aggregate score to a pass requirement for each part. In the same year, the failure rate passed 30 % and the year after even rose to its maximum of more than 40 % in 2014. There may not be single explanation since many variables are at play; however, the pass-both-parts requirement has certainly played a role. In the following years failure rates dropped annually to reach a minimum in 2017.

The professors claimed the workload and difficulty of exams remained stable over time, and students adapted to the new scoring and grading regime after a couple of years. The pass both parts requirement motivated learning to include both algorithmic and conceptual problems. The latter was further addressed in the mid-term exam starting in 2016, exclusively focusing on conceptual themes. A couple of examples from the 2017 mid-term exam may serve to illustrate the tasks: “If a statistical sample is made available to you, describe in a few words which options you would consider for the purpose of choosing a statistical model that could represent the sample in an adequate manner” (Problem 1e, 2017).

--- “Give the properties of a spectrum and make a sketch of how it can possibly look like” (Problem 2e, 2017).

Teaching as problem design draws on an activity-based approach to learning. The interest is in the construction of the learning environment rather than on fine-tuning of lectures based on student satisfaction studies of teaching. The focus is on students’ learning activities rather than what the teacher does, a position also suggested by a seminal research article (Biggs, 1999).

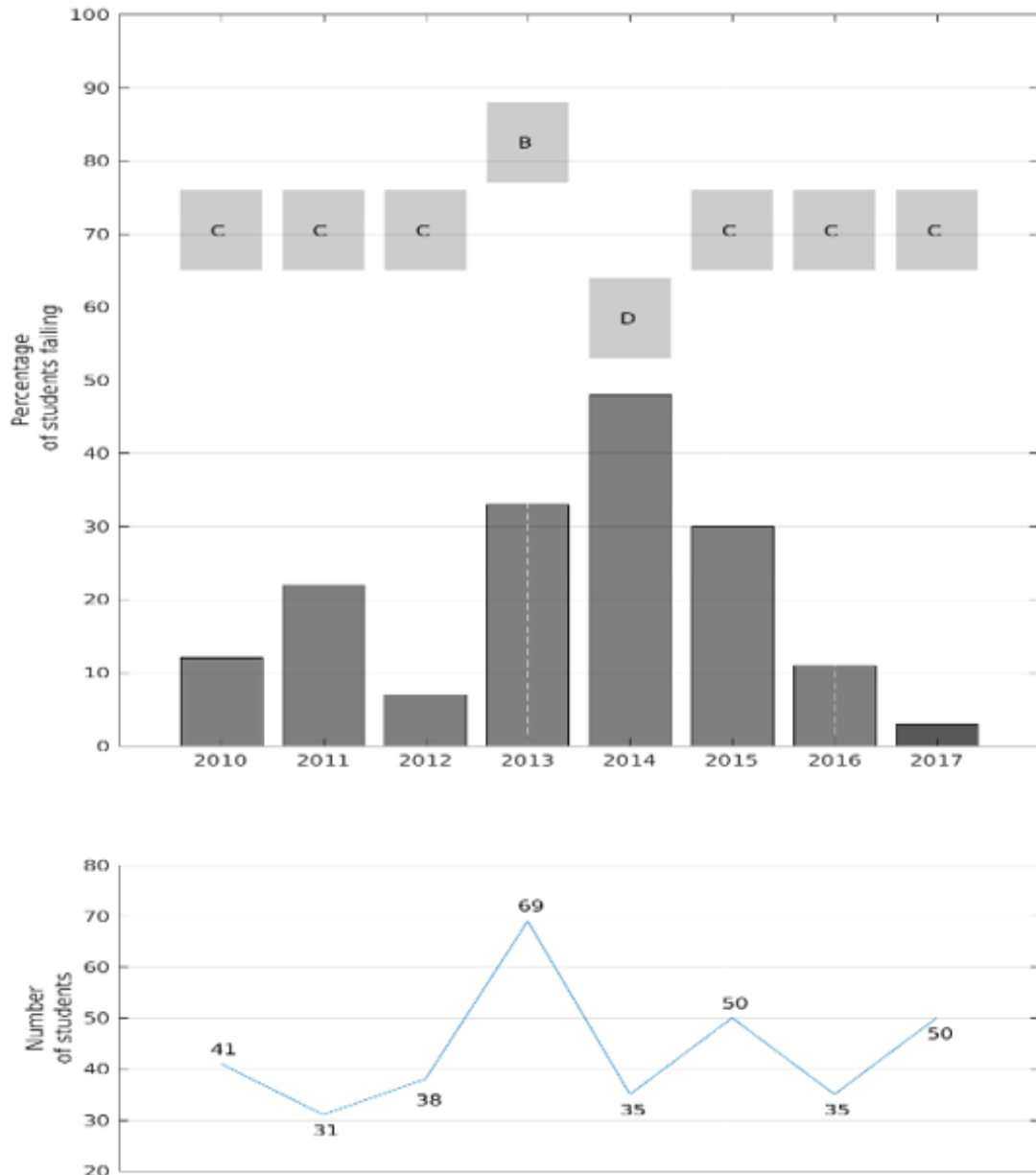


Fig. 1. Number of enrolled students (bottom figure), annual failure rates (middle figure) and mean grades (top figure), 2010-2017

Time pressures and speedy progression frequently make students become strategic in their study orientations (Ramsden, 2003; Snyder, 1971). Professors may therefore benefit from examining effects of past design decisions to prepare future interventions. Figure 2 shows failure rates by year specified on two different causes, insufficient total score versus withdrawals during the exams.

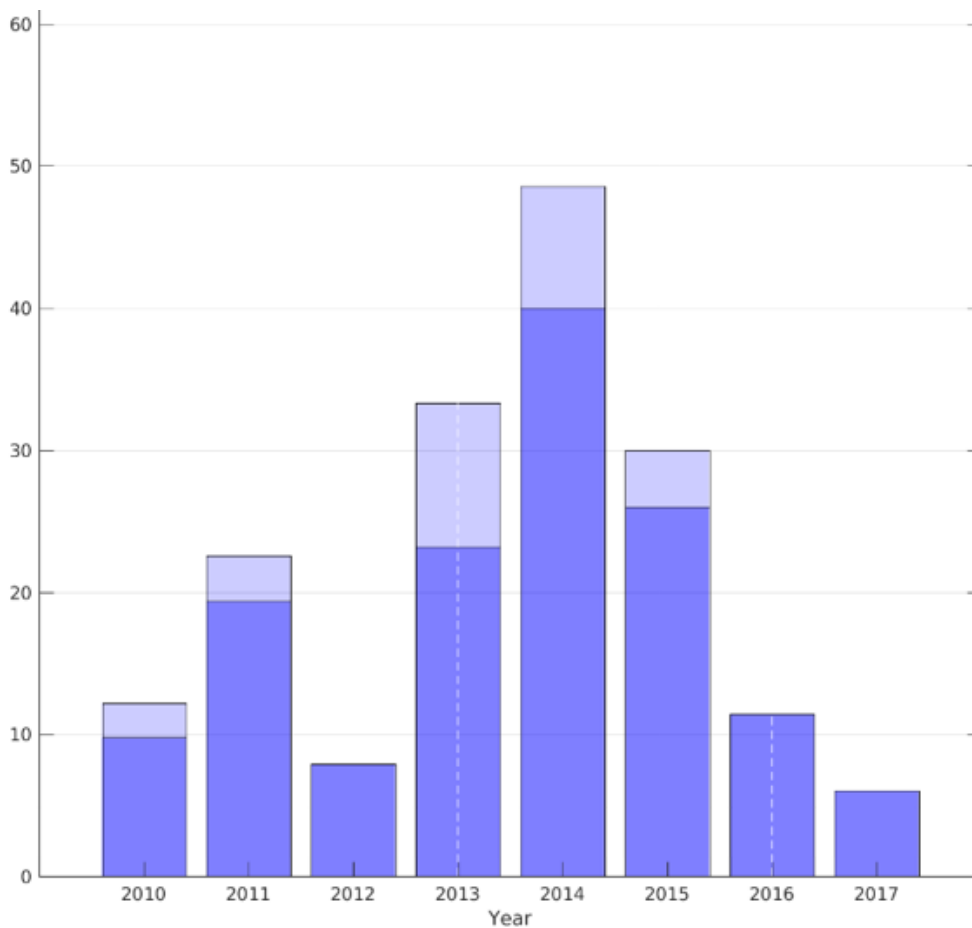


Fig. 2. Blue color indicates insufficient total score (minimum 41 %), while light blue indicates withdrawals during exams. Annual failure rate by year (horizontal axis) and failure rate in percent (vertical axis)

The proportion of withdrawals during the exam was greatest in 2013 when both Part 1 and Part 2 had to be passed to obtain an overall pass. The percentage of failures was sharply reduced from the peak year 2014 to 2017. In the period 2010 through 2012, the failure rate was also relatively low; however, calculated by using a different algorithm compared with the period from 2013 through 2017 when a pass presupposed minimum scores in both parts.

Figure 3 illustrates the relative distribution of the three poorest grades D, E and F. The proportion of failures is relatively low in the three years from 2010 to 2012 with the obvious reason that candidates could pass the course even with a very low score in one of the two parts. A candidate who received a score of 80% in Part 1 and 20% in Part 2 would end up with an average score of 50% and the grade E. When the pass-both-parts regime was adopted in 2013, failure rates increased dramatically compared with previous years, and the proportion of failures reached its peak level in 2014, one year after the new pass requirement was adopted. However, failure rates were sharply reduced in the years thereafter. The authors speculate that students by then benefitted from the revised learning materials and had adapted to the pass-both-parts requirement. The mid-term introduced in 2016 may also have helped to further reduce the failure rate. The proportion of the grade E is lower in all years from 2015 through 2017 compared with the period from 2010 through 2012.

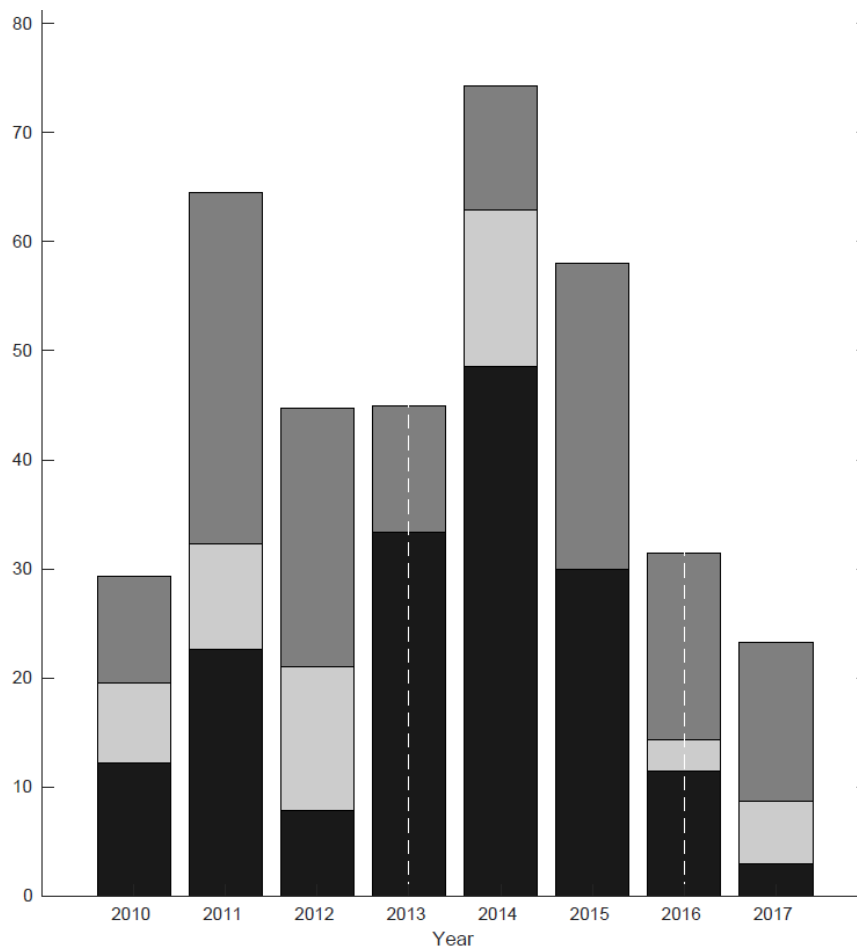


Fig. 3. Distribution of the poorest grades D, E & F from 2010 through 2017. Grey color on top means D, light grey in the middle means E and black color at the bottom means F(ail). Total percentage of the grades D, E & F on the vertical axis

By comparing distributions in Figure 4, an increase in the proportion of the best grades A, B and C is observed, from about 25% in 2014 to almost 80% in 2017. It appears that the new pass requirements yielded a negative effect in the short term, and a more positive effect in the longer term. The relatively low failure rates in 2010 through 2012 may partially be ascribed to pass requirements based on mean scores during those years. Grades assigned from 2013 onwards may be claimed to meet demands of content validity and reliability to a greater extent compared with previous years. Larger parts of the curriculum were incorporated, particularly those contributing to enhanced conceptual understanding.

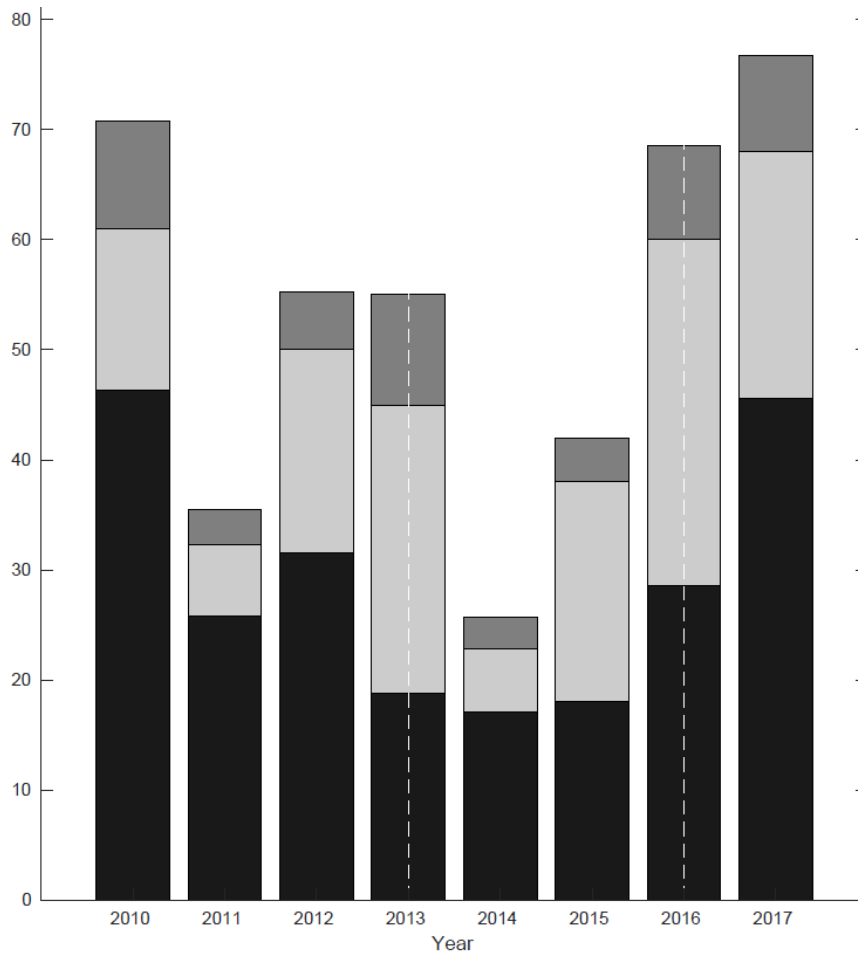


Fig. 4. Distribution of the best grades A, B & C from 2010 through 2017. Grey color on top means A, light grey color in the middle means B and black color at the bottom means C. Total percentage of the grades A, B & C on the vertical axis

Figure 5 illustrates the evolution of mean scores in Part 1 and Part 2 from 2011 through 2017. The “shock” caused by the introduction of the pass-both-parts regime is clearly demonstrated by lower average scores in both parts in 2014. From then onwards mean scores increased and converged, presumably caused by the more even distribution of types of tasks in Part 1 and Part 2, and by improved preparedness to deal with conceptual problems at the final exam.

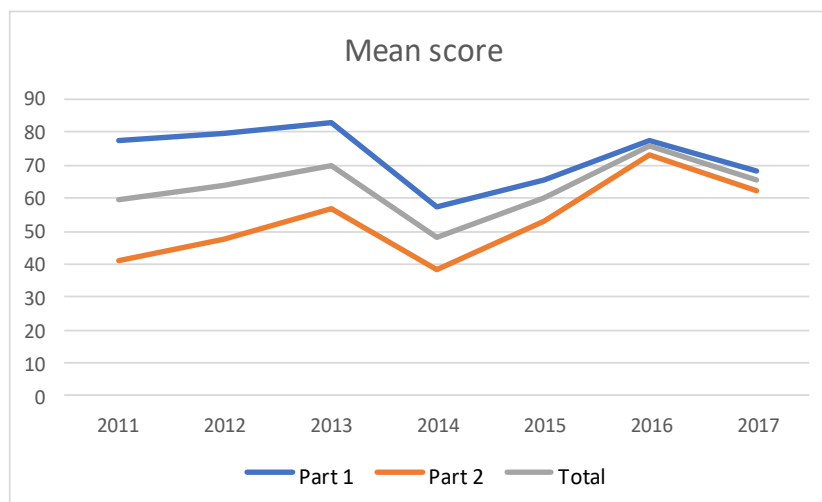


Fig. 5. Mean scores in Part 1 & Part 2, 2011-2017

Shortly after the exams in 2016 and 2017, a survey featuring identical items on a nine-point Likert scale was administered to the students. Approximately half of them responded, and almost all of them took part in the voluntary mid-term exam. Response patterns were largely the same both years with an overwhelming majority claiming that participation in the mid-term had boosted their efforts. However, some assumed the mid-term had negative effects on efforts in parallel courses, while a few others suggested future mid-terms should count only positive, as outlined by a student in the following quote:

To improve learning - make the midterm count positive 30% and hence the students making a great effort through the semester will be rewarded. And in the same manner students doing other things and learning through hard work at the end of the semester will not be punished. I did a good exam, but the only thing an employer will see is my C in the course due to my terrible midterm, because I had other priorities in the middle of the semester. (Student, 2016)

The end-of-term exam succeeded to achieve an even balance between conceptual and algorithmic problems; however, there was no unanimous support among students for the inclusion of more conceptual items. Overall, students rated the course as better in comparison with other courses; however, suggestions for improvements were still numerous:

Try to have physical or real-life examples to connect the theory to something practical. The course is highly theoretical, and it has to be, but I think it will be good for the students to be given some real-life examples in order to get a better understanding of the subject. (Student, 2016).

Part 1) I think the explanation of Monte Carlo simulation was too brief. I think we should have learned about this more in detail. Too much focus on very basic probability calculations during the first 2 lectures (the students know most of this from before). I think more theoretical questions should have been included in the exercises and less calculations in excel. (Student, 2016).

Part 2) Time on the different concepts was well spent. I think some of the calculations was a bit too long to go through in class, but the level of the class was in general very suitable after having had "Marine Dynamics" previously. I would have liked more time to understand the concept of frequency response function and impulse response function. (Student, 2016)

6 DISCUSSION

This study starts out by addressing issues of learning rather than teaching, indicating a distinction between means and ends in education. Historically, in the course under scrutiny the focus had been on satisfaction studies of teaching, without attempts being made on theorizing relationships between teaching and learning. Consequently, remedial measures aimed to enhance satisfaction of teaching rather than exploring and theorizing causes of learning deficiencies. The conception of teaching was restricted to interactive forms of teaching, while pre-active and post-active forms were omitted. The construction of survey instruments is therefore an issue, since conceptual understandings of key terms and their relationships inform and determine the nature and use of collected data.

This study utilizes and extended definition of teaching as anything impacting learning. Quality enhancement took the form of grounded steps based on an explanatory understanding of repeated patterns of poor learning, calling for a rigorous search for potential causes. Addressing task and problem design in exercises and at exams appeared as a sensible approach since students made rational choices of what exactly paid off at the end-of-term exams. Compared with fine-tuning of lectures problem design proved far more efficacious; however, still not used as part of the university's induction courses for new academic staff. Teaching as design is an activity-based approach, emphasizing doing as a means towards ends. Most learning takes place without professional support, and consequently students need to engage with tasks that are tailor-made and conducive to intended learning outcomes.

Algorithmic problems can be solved by memorized sets of procedures, while a conceptual problem requires the student to work from an understanding of a concept where no memorized procedure is likely to be known (Cracolice, Deming, & Ehlert, 2008). Certain types of problems are known to be more conducive to success than others. For example, one study reported far greater success among students solving algorithmic problems compared with open-ended questions (Surif, Ibrahim, & Dalim, 2014). Another study found that many students have "... no choice other than to be algorithmic problem solvers because their reasoning skills are not sufficiently developed" (Cracolice, Deming, & Ehlert, 2008). It

has been theorized that conceptual learning is of a different kind compared with memory-based learning (Rillero, 2016), and that limited ability to cope with concepts might have been caused by an overemphasized focus on algorithmic problem solving (Igaz & Proksa, 2012).

The construction of targeted learning experiences aligned with assessment helped to shape a learning environment conducive to desired learning. Previously, time and effort were spent on assessing learning outcomes with grades as the sole product. This study demonstrates an extended use of performance data to the benefit of learning. Constructive alignment was used as the overarching principle in weaving together measures into an integrated whole. Ownership to, and participation in the project, empowered and motivated the professors for continuous improvement. Generally, learning design competence would be one way to make educational interventions more focused and productive as an alternative to, or in combination with, the widespread use of satisfaction studies. The application of action research enables the combination of practical interventions and evidence of outcomes, and disparities between the two serve to prompt the continuous theorizing of causes ensued by grounded action.

7 CONCLUSION

High failure rates and poor academic performance is common in science and engineering education; however, in many instances such practices prevail despite sustained efforts to remedy the situation. In this study, the professors received favorable responses on their teaching, while academic achievement generally remained below expectations. The lesson learned can be summarized in a few sentences: The focus in educational development needs to switch from teaching to learning, and the concept of teaching should be expanded to include a broader range of approaches of relevance to learning. This implies a clear distinction between means and ends, and the application of an explanatory approach to deficiencies enables a more targeted use of interventions. Still, too many variables are at play to make predictions certain, and the application of trial and error is always a useful aid. A systems approach does not imply simple, formula-like solutions; however, in-depth understanding of key educational concepts may empower faculty to become more competent and self-reliant practitioners.

REFERENCES

- Barr, R. B., Tagg, J. (1995). From Teaching to Learning: A New Paradigm for Undergraduate Education, *Change*, Vol. 27, No. 6, pp. 12-25.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347-364. Retrieved from <http://dx.doi.org/10.1007/BF00138871>. doi:10.1007/bf00138871
- Biggs, J. (1999), What the Student Does: teaching for enhanced learning. *Higher Education Research & Development*, Vol. 18, No. 1, pp. 57-75.
- Cracolice, M. S., Deming, J. C., Ehlert, B. (2008), Concept Learning versus Problem Solving: A Cognitive Difference. *Journal of Chemical Education*, Vol. 85, No. 6, pp. 873-878.
- Goodyear, P. (2015), Teaching as design. *Herdsa Review of Higher Education*, Vol. 2, pp. 27-50.
- Igaz, C., Proksa, M. (2012), Conceptual Questions and Lack of Formal Reasoning: Are They Mutually Exclusive? *Journal of Chemical Education*, Vol. 89, pp. 1243-1248.
- Land, R., Cousin, G., Meyer, J. H. F., Davies, P. (2005), Threshold concepts and troublesome knowledge (3)*: implications for course design and evaluation. In C. Rust (Ed.), *Improving Student Learning Diversity and Inclusivity*. Oxford: Oxford Centre for Staff and Learning Development.
- Marton, F., Säljö, R. (1976a), On Qualitative Differences in Learning — 1: Outcome and Process. *British Journal of Educational Psychology*, Vol. 46, pp. 4-11.
- Marton, F., & Säljö, R. (1976b). On Qualitative Differences in Learning — 2: Outcome as a function of the learner's conception of the task *British Journal of Educational Psychology*, 46, 115-127.
- Ramsden, P. (2003). *Learning to teach in higher education*. London: Routledge Falmer.
- Rillero, P. (2016), Deep Conceptual Learning in Science and Mathematics: Perspectives of Teachers and Administrators. *Electronic Journal of Science and Education*, Vol. 20, No.2, pp. 14-31.
- Rust, C. (2002), The impact of assessment on student learning. *Active Learning in Higher Education*, Vol 3, No. 3, pp. 145-158.
- Simmons, O., Gregory, T. (2003), Grounded Action: Achieving Optimal and Sustainable Change. *Forum: Qualitative Social Research*, Vol. 4, No. 3.
- Snyder, B. R. (1971), *The hidden curriculum*. New York: Knopf.
- Surif, J., Ibrahim, N. H., Dalim, S. F. (2014). Problem Solving: Algorithms and Conceptual and Open-ended Problems in Chemistry. *Procedia - Social and Behavioral Sciences*, Vol. 116 (Supplement C), pp. 4955-4963. doi:<https://doi.org/10.1016/j.sbspro.2014.01.1055>