Dansk Universitetspædagogisk Tidsskrift

Tema Nye perspektiver på evalueringsformer i universitetspædagogik

Årgang 12 nr. 23 / 2017

Titel **Continuous assessment in higher education in Denmark: Early experiences from two science courses**

Forfatter Ole Eggers Bjælde, Tove Hedegaard Jørgensen og Annika Büchert Lindberg

Sidetal 1-19

Udgivet af Dansk Universitetspædagogisk Netværk, DUN

URL http://dun-net.dk/

Continuous assessment in higher education in Denmark: Early experiences from two science courses

Ole Eggers Bjælde, Educational developer, astrophysicist and special consultant, PhD, Science and Technology Learning Lab, Faculty of Science and Technology, Aarhus University

Tove Hedegaard Jørgensen, Associate Professor in Evolutionary Biology Bioscience, Faculty of Science and Technology, Aarhus University

Annika Büchert Lindberg, Educational developer, tropical biologist and special consultant, PhD, Science and Technology Learning lab, Faculty of Science and Technology, Aarhus University

Research article, peer reviewed

Designing fair and efficient ways of assessing student learning is a challenge to most teachers in higher education. It is possible that multiple graded, low-stake activities during the teaching period can either replace or supplement end-of-semester exams to measure student performance. Such a shift to continuous assessment has the potential not only to increase efficiency but, importantly, also enhance student learning. Continuous assessment is used widely internationally and now (since 2016) also allowed at Danish Universities. Here we review the advantages and disadvantages of this assessment format and report on its first use in two science courses at Aarhus University. We include a detailed description of the graded tasks and activities used in the two courses. By comparing student performance in continuous assessments with that of a traditional end-of-semester exam we are able to highlight some challenges and provide recommendations for the future use of this assessment format at Danish universities.

Introduction

The Ministerial Order for Examination (30/06/2016) now allows the use of continuous assessment at Danish Universities. It is an assessment format that has the potential to change student study behaviours while it also offers the opportunity to provide more feedback and to improve the alignment between teaching and exams. One expectation is that effective use of continuous assessment can boost student completion rates and reduce drop-out rates through enhanced learning and the avoidance of single high-stakes exams. Here we give an introduction to the potential uses, advantages and disadvantages of continuous assessment. We furthermore

describe and discuss some first experiences of using this assessment format in two undergraduate courses at a Danish university.

Term	Additional terms	Definition	In Danish
Assessment		'graded and non-graded tasks, undertaken by an enrolled student as part of their formal study, where the learner's performance is judged by others (teachers or peers)' (Bearman et al. 2016, p. 547).	Bedømmelse/ udprøvning
Examination/exam	End-of-semester assessment/final assessment/final exam	'Assessment undertaken in strict formal and invigilated time-constrained conditions' (Bridges et al. 2002, p. 36). Is graded.	Eksamen
Continuous assessment	Coursework/ curriculum integrated assessment/ embedded assessment	Assessments occur as graded tasks or activities (written assignments, tests, small oral presentations and similar) distributed throughout the course.	Løbende bedømmelse
Formative assessment		General term for non-graded assessments that can be distributed throughout the course and provide the opportunity for feedback and feed-forward. Used by teachers and students to adjust teaching and learning activities (Black and Wiliam, 1998).	Formativ bedømmelse
Summative assessment		General term for graded assessments that provide information about the level of student performance. These assessments can be distributed throughout the course. (Trotter 2006)	Summativ bedømmelse
Evaluation	Course evaluation	Student evaluation of the teaching/instruction during the course.	Evaluering

Table 1: Definitions of terms are based on the literature where possible and on own wording in the remaining cases.

Assessment and learning

Assessment plays an important role in student learning and is perhaps the most important factor for student motivation and engagement (Ramsden, 2003; Brown et al., 1997). In this paper, we define assessments as 'graded and non-graded tasks, undertaken by an enrolled student as part of their formal study, where the learner's performance is judged by others (teachers or peers)' (Bearman et al., 2016, p. 547, see also Table 1 for definitions of terms). Assessment has three main functions 1) to assign grades that judge the quality of student achievements, 2) to provide evidence or certification to external partners and 3) to support student learning (Carless, 2015). Functions one and two are referred to as *assessment of learning* and are well described in university policies on assessment (Boud, 2007).

The traditional time-bound, unseen and written end-of-semester examination serves these functions by striving for reliable and fair assessment with limited possibilities for cheating (Race, 2014). Oral assessments, where the students draw a question to be answered and discussed immediately or after a short preparation time (Ulriksen, 2014), also assess learning and are commonly used in Scandinavia and Germany (Andersen & Tofteskov, 2016). End-of-semester examinations provide limited opportunities for feedback to learners and, in their typical form, reveal little information that might help students improve their understanding. One can argue that this kind of examination does have some formative elements because students can adapt their learning activities to this particular assessment format, e.g. answering questions or solving problems from previous examinations. Still, the main function of end-of-semester examinations is to test whether students meet a given standard (Raaheim, 2016). They become high-stakes because students usually have only one chance to deliver and may therefore promote exam anxiety. The examinations also often lack authenticity in the sense that they rarely mirror real-life tasks or real-life conditions and usually require students to work alone, with limited access to resources and with minimal influence on the assessment task itself.

A particular challenge for those involved in creating assessments is to find a design that facilitates the long-term retention of learning. This is not always the case with traditional, time-bound, end of semester examinations where students often revise intensively before sitting the exam, but find they have forgotten much of what they revised, once the examination is over. The question is whether the use of other assessment formats can help teachers meet some of the challenges posed by final examinations and move the emphasis from control of standards and certification to also include authenticity and emphasis on learning.

Continuous assessment for learning

When assessments occur as graded tasks or activities distributed throughout the course (written assignments, tests, small oral presentations and similar) we refer to them as continuous assessment. Each separate assessment will count towards the final grade and can be regarded as a formative/summative hybrid because it can include increased opportunity for learning (hence the term 'learning-oriented assessment' used by Carless 2007). Low stake summative assessment tasks can engage students throughout the course and define standards against which students can test their understanding formatively, thus helping students to internalise these same standards.

The idea of using continuous assessment in higher education is not new. End-of semester examinations have been supplemented or replaced by continuous assessments in the UK, Australia and New Zealand over the last 40 years (Richardson 2015). Also Universities in the USA have used continuous assessment for decades. For example at Harvard University where a final exam can now (since 2010) only be held by special permission as a supplement to the continuous assessment (Harvard Magazine, 2010). Another example is the University of Western Australia where final high-stakes exams will be removed from timetables in 2018 and replaced by a format where any one assessment task must comprise less than 70 per cent of the final grade, including a potential final exam (University of Western Australia, 2015). In a Danish context, this form of assessment has only recently become available to teachers in higher education.

Advantages and uses of continuous assessment

Assigning grades and certification is an important purpose of assessment because it affects the future careers of students (Boud and Falchikov, 2007a). However, the potential use of assessment *for* learning and not just *of* learning is increasingly accepted in higher education (Brown, 2005; Boud & Falchikov, 2007a). Without dismissing the certification aspect of assessment we focus in this section on the learning-oriented aspects of continuous assessment and summarise the possible advantages.

Boosting student motivation with continuous assessment

Students' engagement in assessment activities is influenced by their perception of assessment purpose (Carless 2015). Making the assessment summative can therefore be an important incentive for students to perform at their best (Carless, 2015). If activities are instead voluntary or serve as prerequisites for an end-of-semester examination, students are less likely to put real effort into the activities. If, however, feedback consists of formative feedback as well as a summative grade, this can potentially increase student motivation for engagement in the curriculum throughout

the course and avoid 'last minute cramming' before the final examination (Trotter, 2006; Gibbs and Lucas, 1997).

Using continuous assessment to strengthen practice and the effectiveness of feedback

One example of the learning-enhancing aspects of continuous assessment is the opportunity to practise skills that can be improved when the students are provided with (timely) feedback and the opportunity to follow-up or act upon the feedback (Bearman et al, 2014). This becomes particularly powerful when activities are also graded, because the combination of grade and feedback will hold more information than either one of them alone. Moreover, due to a lack of time constraints, students are more likely to produce work of academic excellence (Bassey, 1971; Richardson, 2015). Assessments can then improve knowledge and understanding, as well as provide practice of specific skills like writing, presenting, problem solving, handling equipment, etc. When used in this way, continuous assessment offers a way of integrating student learning progress into the assessment so that attention is not only on the end result but also on the learning process (Ramsden, 2003; Dochy et al., 2007). This idea of rewarding increased effort and persistence rather than focusing on actual performance is in line with recommendations from studies of metacognition (Schraw, 1998 and references therein). A multi-step assessment activity with feedback at intermediate stages could be one way of achieving this (see 'Assignments' in box 1 for an example). This format also addresses the challenge of ensuring that students use feedback constructively in later assessments. As a final remark on feedback, continuous assessment will also inform teachers about student progress and learning and thereby help expose areas in teaching where adjustments may be needed.

Mirroring real-life tasks with continuous assessment

An interesting aspect of continuous assessment is that it offers a way to test competencies that can be hard to assess in a traditional final exam. This includes competencies such as the ability to collaborate with peers, creative thinking and innovation skills (Bjælde and Najbjerg, 2017). Assessment tasks can therefore be more authentic because the working process in the assessment can resemble more closely the study process students are used to in their course of study. Furthermore, authentic (and graded) tasks mirroring a future professional life (e.g. law students identifying legal issues reported in news media or medical students treating patients) may be highly motivating for students (Glofcheski, 2017). However, setting assessment criteria for such authentic tasks and explaining to students exactly how they will be assessed can be challenging (Bridges et al. 2017).

Helping students to become self-reflective learners

The ability to judge the quality of your own work is a required competence in students' future professions (Boud and Falchikov, 2007b). Engaging in meaningful continuous assessment activities that focus directly on the application or development of assessment criteria is one way to practise this (see example in Box 1). Such activities may generally strengthen the student's beliefs in their own abilities as a learner (Shields, 2015). This effect can be further increased by letting students use criteria to assess the quality of their own work (without marking) (McDonald and Boud, 2003; Andrade & Du, 2007). Acquiring a detailed understanding of the standards within their discipline will also help students prepare for a potential final examination.

Assessing large cohorts with continuous assessment

Many continuous assessment activities are well-suited as online activities in a Learning Management System. It can therefore make continuous assessment feasible even for larger cohorts because it can reduce marking time, provide opportunities for automated feedback and support student engagement with feedback (Bennett et al., 2016). A few common examples are multiple choice questions or short essays with word restrictions and rubrics for transparent and fast marking (see also Box 1 for examples). However, it is important to consider how technology and pedagogy can be combined to improve assessment for learning and not just of learning (Dawson & Henderson, 2017). Technology-supported assessment poses a risk of focusing more on efficiencies in assessment (e.g. reduction of marking time) and not on assessment for learning through innovative assessment tasks.

Exam anxiety

Assessments (and examinations in particular) are potentially very stressful to students (Falchikov and Boud, 2007c). It is a highly undesirable situation because assessments are designed to focus on student achievement of learning outcomes and not on their ability to handle stress. It is possible that low-stake assessments providing timely feedback to students can be experienced as less stressful for the majority of students and increase their confidence (Shields, 2015). Continuous assessment with many low-stake assessments may be particularly useful when helping first year students to understand expectations in higher education.

Challenges with continuous assessment

Replacing one assessment practise with another obviously requires a time investment from teachers and may incur additional costs. Continuous assessments combining both grading and feedback may be particularly costly to design and implement (Hernandez, 2012; Carless, 2015). It is for example time consuming to design

assessment tasks and feedback so that students use the feedback to support their learning and not as simple explanation for the grade (Glover & Brown, 2006; Hernandez, 2012). Box 1 provides examples of multi-step assessments designed to do exactly this. Another challenge is to ensure that each assessment task is viewed as part of a coherent curriculum and not as an isolated piece of learning, for example by letting students compare and contrast or link different subjects or concepts from the curriculum to each other. In some cases a supplementary final examination can also be used to bring all the pieces of the curriculum together again.

Using continuous assessment comes with the risk that students may feel as if they are constantly assessed. So, instead of reducing anxiety, continuous assessment may actually make students feel anxious for more of their study time. In order to avoid this risk, the aim should be to strike a meaningful balance between complexity of activities, the number of activities and the available time. In general, students are willing to invest a considerable amount of time and effort when learning activities are perceived as meaningful and involve a substantial degree of challenge (Marsh, 2001; Trotter, 2006; Raaheim, 2016).

A final concern when engaging in continuous assessment is the issue of cheating and plagiarism. Because students are not assessed under strictly controlled conditions they will have access to all available resources and aids. A few Norwegian studies show that frequent assignments, too many assessments and pressure for good grades are among the main reasons why some students cheat (Raaheim, 2016, and references therein). Ignorance on what cheating and plagiarism is and the simple possibility to cheat are additional reasons (Park, 2003; Raaheim, 2016 and references therein). Hence, cheating is a real concern, and continuous assessment activities have to be designed carefully to avoid it. But, it can also be argued, that cheating is, and always has been, a problem for end-of-semester examinations too. Continuous assessment can even include activities that develop student understanding about cheating and plagiarism. Additionally, continuous assessment can focus on students' reflections and responses to various sources of information and less on checking if students have acquired specific knowledge (Raaheim, 2016). The use of digital platforms for assessment can also help as many offer plagiarism checking and can generate unique questions and tasks for each student, making it harder to cheat.

Continuous assessment in the Danish context

The Ministerial Order from 30.06.2016 (Ministerial Order, 2016, p. 2) now offers universities in Denmark the possibility of using continuous assessment:

> *In the academic regulations, the university may also stipulate that the assessment of coursework in the form of written papers and oral presentations etc. must be in-*

> *cluded in the determination of the mark together with the final exam in a course or course element.*

Note the flexibility in assessment format that this wording allows: '*written papers, oral presentations, etc*' can be included in the calculation of the final grade. The Ministerial Order further underlines that students should know exactly how their grades are calculated (Ministerial Order, 2016, p. 2):

> *It must be stated in the rules, if any, how the assessment of the written papers and oral presentations etc. should be included in the overall assessment of the course or course element.*

A major issue in higher education in Denmark is the call for more feedback. This was identified as the most important action point for Danish universities in a survey of 76,000 students and 43,000 new graduates in 2017 (Uddannelses- og Forskningsministeriet, 2017). Due to the reinforcing effect of combining feedback with continuous assessment, allowing and encouraging the use of continuous assessment therefore appears timely. Currently we have only a few documented experiences from the use of continuous assessment at Danish universities (see Christensen (2016) for an example combining continuous assessment and agile feedback). Below we report on our first use of this assessment format in two undergraduate courses in Physics and Biology at Aarhus University and discuss their outcome in relation to expectations from the literature. Note that both course organisers are among the authors of this paper.

Early experiences of using continuous assessment at Aarhus University

Astrophysics

Astrophysics is a 5 ECTS mandatory course in the first semester of the physics undergraduate programme with 100-150 students per year. The course serves as an introduction to the field of astrophysics and covers a broad curriculum that includes many different topics. During each week of the semester, the course has three hours of lectures, three hours of exercises/tutorials and a substantial online component corresponding to roughly 25 per cent of the course work. Assessment includes a continuous component (online) and a final exam (on-site). The continuous component consists of reading quizzes, assignments and communication exercises, organised in a weekly structure. Each element is described in more detail in Box 1. The final exam is a three-hour written exam with an emphasis on problem solving.

In the 2016 edition of Astrophysics, the continuous assessment activities were the following (percentage of final grade in parenthesis):

- Reading quizzes (8 %)
  - When: Week 1-7
  - What: Multiple choice questions, ordering questions, matching questions, etc
  - Feedback: Students can answer the questions as many times as they like, only the last attempt counts. After each attempt, students get automated feedback pointing towards the correct answers in the book.
- Criteria exercise (2 %)
  - When: Week 1
  - What: Students rank four different written answers to a problem and give criteria and explanations for the ranking. The written answers are anonymised student answers from a previous year.
  - Feedback: Students perform this activity in a group and collective feedback is given to all groups. In addition, all student criteria are collected and merged into a list of assessment criteria that are used to score assignments in subsequent weeks.
- Assignments (14 %)
  - When: Week 2, 3, 4
  - What: Problems from a previous final exam
  - Feedback: Students get feedback from a teaching assistant. They can then resubmit the assignment taking into account the received feedback to obtain a better score.
- Design a multiple choice question (6 %)
  - When: Week 5-7
  - What: Students create their own multiple choice question in the online system PeerWise (Denny et al., 2008) including plausible wrong answers and an explanation for the correct answer. The question is then uploaded to a common question pool in PeerWise.
  - Feedback: Students receive points by creating at least one question and answering and rating 20 questions from the question pool. Bonus points are awarded to students with high-rated questions and with many badges (assigned automatically by the system).
- Communication exercise (20 %)
  - When: Week 5-7
  - What: In groups students communicate a topic from the curriculum to a selected audience. Assessment criteria include subject knowledge and coherent communication but also innovation and multimodality (more than one mode of communication). All criteria are known to students beforehand. The format of presenting can be chosen freely by students, but they are encouraged to create a product which will give actual value for their selected audience.
  - Feedback: Students' products are graded with a rubric and short, targeted feedback is given after submission. Read more about the communication exercise in Author et al. 2017.

Box 1: Continuous assessment activities in Astrophysics in 2016.

Continuous assessment was introduced in the course in 2014 (by dispensation) to motivate students to work in a structured manner throughout the course, to avoid a single high-stakes exam at the end and to support student learning by providing timely feedback on graded learning activities. The structure and cadence of activities in the course were designed using a learning design model (Godsk, 2013; Bjælde et

al., 2015). The activities in the continuous assessment all take place online, and students are allowed the flexibility to do the activities whenever they want (before deadline) and with whom they want. The content in the continuous assessment activities in a given week mirrors the content covered in lectures and exercises in that week. In practically all continuous assessment activities, students would benefit directly from working in a group and discussing problems with other group members. For this reason in particular, it is expected that students on average perform well in the continuous assessment. In addition, the structured and persistent work required to do well in the continuous assessment is expected to boost the grade point average and lower fail-rates.

Overall grades in Astrophysics for the years 2013-2016 as well as grades in continuous assessment vs. final exam from 2016 are shown in Fig. 1A, 1C and Table 1 (cohort sizes given in Table 1). There are several noteworthy trends in the grades; first of all the introduction of continuous assessment has lowered the fail-rates and increased the grade point average. Moreover, students' performances in the continuous assessment activities are significantly better than in the final written exam, as expected. We defer a further discussion of these numbers to the next section.

Towards the end of the teaching period each year, a student evaluation survey is completed gauging, among other things, students' opinion on continuous assessment. In 2015 and 2016 students were directly asked how many per cent continuous assessment should contribute to the overall grade. To this, students responded 43 per cent in 2016 (N=87), and 23 per cent in 2015 (N=91), in both cases with a large variation. It is interesting to note, that in 2015 continuous assessment counted for 25 per cent of the total mark and in 2016 this was increased to 50 per cent more or less matching students' preferences. Students were also asked directly whether they supported the use of graded continuous assessment activities to which students in 2016 responded: yes (79.5 per cent), no (11.4 per cent) and don't know (9.1 per cent). The average student evidently supports the use of continuous assessment. A third interesting number from the student evaluation survey is the perceived number of hours spent on astrophysics per week. The 2016 number is not reliable due to the lack of respondents, but the numbers from the previous years were (number of responses in parenthesis): 2015: 11.1 hours (N=91); 2014: 9.8 hours (N=89); 2013: 12.0 hours (N=40). Note that continuous assessment was introduced in 2014. Students' perceived workload does not seem to increase with continuous assessment when looking at the entire ensemble of students.

Evolutionary Biology

Evolutionary Biology is a 5 ECTS mandatory course in the fourth semester of the Biology degree. Teaching is delivered through four lectures and two hours of small group teaching per week for seven weeks and it gives 100 - 130 students a first in-

troduction to the field. Students are asked to prepare for in-class activities through directed reading, watching of webcasts, solving quizzes and working on larger analytical problems. The latter forms the basis of activities during small group teaching. The course was previously assessed in one final four-hour written exam consisting of analytical problems and multiple-choice questions in exactly the same format as students had met them during teaching. In 2016 we moved (by dispensation) to a combination of continuous assessment (online; accounts for 25% of the course mark) and a final three-hour written exam (on site; accounts for 75% of the mark) but kept the existing format for test questions. Our aims were threefold: 1) to avoid a single high-stakes exam, 2) provide opportunities for feedback to students throughout the teaching period and 3) enhance student learning through increased engagement in both out-of-class and in-class activities. The continuous assessments were designed to reward investment in preparation and active participation in the small group teaching in particular. Each weekly assignment tests the students' understanding of exactly the same concepts as covered during teaching that week and often is based on the same examples and data as used in small group teaching although the questions vary. It was expected that a high learning outcome from the small group teaching would lead to a high mark in the weekly assignments. These assignments are made available online midweek when all teaching for the week has finished. The students then have five days to complete the assignment with the possibility to resubmit and also to collaborate with peers.

One expectation was that the new assessment format would increase the average grade and/or lower the number of students failing the course. Our first experience with the format does not completely match these expectations. While the average grade for the continuous assessments in 2016 was high, average grades for the final exam were lower and left the overall mark for the cohort in 2016 unchanged from previous years (Table 1, Fig. 1B). One observation was that the number of students obtaining a grade of zero or two in the final exam was high in 2016 compared to previous years despite many of these students performing well in the continuous assessment (Fig. 1D).

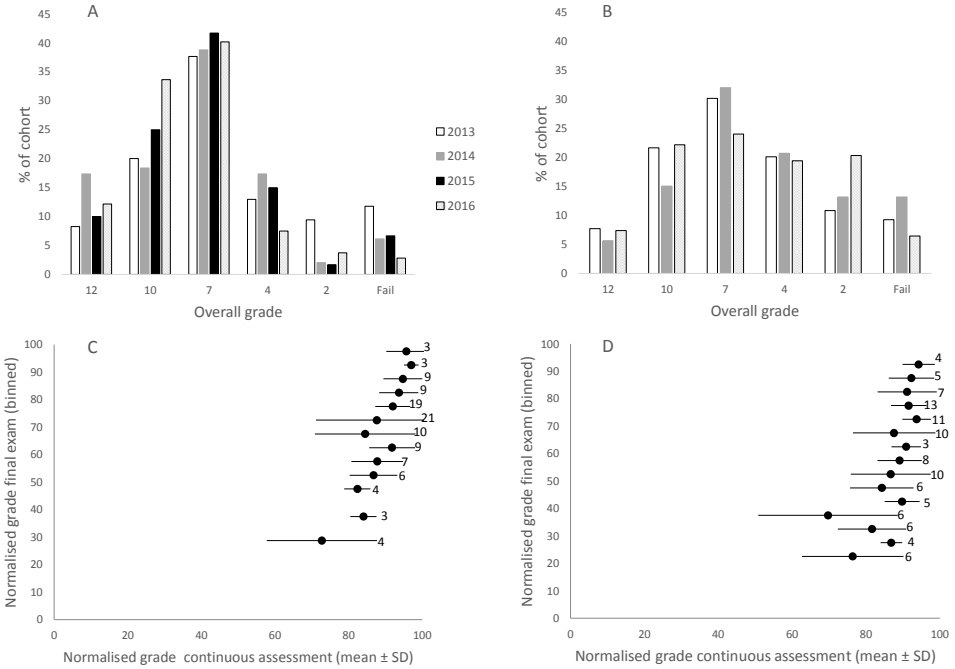

Figure 1: Overall grades awarded to students in Astrophysics (A) and Evolutionary Biology (B) and the association between continuous assessment grades and the final exam grade in 2016 for Astrophysics (C) and Evolutionary Biology (D). Astrophysics used continuous assessment in 2014, 2015 and 2016 and Evolutionary Biology in 2016. The 2015 cohort in Evolutionary Biology is not reported because irregularities (plagiarism) significantly affected the distribution of course marks. Data in panel C and D is binned according to final grade (5% intervals) and the associated continuous assessment grades are reported as means ± 1 SD for full anonymity. Figures next to error bars give student numbers in each bin.

	Overall	**Continuous assessment**	**Final exam**	**N**
Astrophysics				
2013	6.3 (± 3.7)			85
2014	7.4 (± 3.3)	10.0 (± 3.2)	6.7 (± 3.5)	98
2015	7.3 (± 3.1)	7.8 (± 3.7)	6.9 (± 3.2)	120
2016	8.0 (± 2.8)	10.4 (± 2.4)	6.0 (± 3.1)	107

	Overall	Continuous assessment	Final exam	N
Evolutionary Biology				
2013	6.2 (± 3.7)			129
2014	5.5 (± 3.6)			104
2016	6.0 (± 3.6)	10.5 (± 2.4)	4.9 (± 3.7)	104

Table 1: Grade means (± 1 SD) in Astrophysics and Evolutionary Biology in 2013 - 2016. Astrophysics used continuous assessment in 2014, 2015 and 2016 and Evolutionary Biology in 2016. The 2015 cohort in Evolutionary Biology is not reported because irregularities (plagiarism) significantly affected the distribution of course marks. Note that students were not given separate grades for the continuous assessment and final exam, but only an overall grade. The grades shown here for continuous assessment and final exam were calculated using the same algorithm as used to calculate the overall grade.

Discussion of the early experiences

The grades awarded in our first use of continuous assessment in Astrophysics and Evolutionary Biology show that students in both courses perform very well in continuous assessment activities. This is not surprising, and similar results have been obtained across many British universities (Yorke, Bridges and Woolf, 2000; Bridges, 2002; Simonite, 2003). The interpretation here is that increased performance is explained by students having control of the effort invested in continuous assessment activities, the availability of information, the availability of relatively unlimited time in continuous assessment and collaborative working (Yorke, Bridges and Woolf, 2000; Bridges, 2002). In our case, a reasonable suggestion is that many students have benefitted from collaborating in groups and from investing the time required to do well. However, our (limited) data on student behaviour do not immediately support (or dismiss) this hypothesis as exemplified by the reporting of a more or less unchanged perceived workload by students in the Astrophysics course. This does of course not change the fact that a good performance in the continuous assessment is a highly desirable result in itself.

Grades awarded in continuous assessment in the latest installment of both courses show a small variation, whereas marks in the final exam showed a larger variation, demonstrated by the larger standard deviation. A UK study reports the same ten-

dency (Simonite, 2003). This does not necessarily pose a problem, as it may simply show that students have learned from collaborative learning on continuous assessments, thus evening out the grade distribution. More research is needed to clarify this issue.

A reasonable assumption would be that students who perform well in the continuous assessment would be better prepared for the final exam, under the assumption that similar competencies are required in the continuous assessment and final exam. A close association between final exam grades and completion of all continuous assessment activities was reported in a study from the University of Maastricht in the Netherlands (Gijbels et al., 2005). However, the data from the two courses presented in this paper show no close association between the performance in the continuous assessment and in the final exam. This is most visible from the plots in panel C and D of Fig. 1. In both courses, many students, who obtain very high scores in the continuous assessment, obtain below 50 per cent of the possible points in the final exam. One possibility is that a good performance in the continuous assessment activities might lull students into a false sense of security, although there is no data to back-up this suggestion at this stage. The same patterns observed at Aarhus University have also been noted at some British Universities and the interpretation here is that continuous assessments and final examinations do not test the same competencies (Yorke, Bridges and Woolf, 2000; Bridges, 2002). It is for example argued that final examinations will test students' ability to organise knowledge under pressure, while this is less important during continuous assessment (Yorke, Bridges and Woolf, 2000; Bridges, 2002). Additionally, final exams (in the British design) rely heavily on memory since all preparation has to take place before the final examination. The examination conditions, it is argued, simply prevent the students from delivering their best work (Yorke, Bridges and Woolf, 2000; Bridges, 2002). In our case, similar reasons are possible, however, there is also the option that a good performance in the continuous assessments may have lowered the motivation for revision and exam preparation in some students. A different explanation could be that they represent a group of students who underperform due to test anxiety in the final exam and that they simply benefit from continuous assessment where this anxiety is less pronounced (Falchikov and Boud, 2007c; Shields, 2015). The unexplained patterns call for a closer investigation of student motivations and behaviours through focused interviews. At the present stage we can conclude that the activities and tasks used in continuous assessment activities alone or in the final exam alone in the two courses may not be sufficient to accurately assess the competencies and skills of different students.

Assessment activities, in both cases presented, were designed to avoid a single high-stakes final exam, introducing more feedback to students during the semester and to generally strengthen student learning. As teachers and course organisers we

gained important opportunities to judge the quality of student achievements throughout the course, which allowed us to adjust our own teaching and instruct teaching assistants accordingly. The time spent on design, preparation and feedback was not recorded but we judge this to be somewhat higher than before continuous assessment was introduced. Students in both courses have received more feedback compared to students in the years before continuous assessment was introduced and they have been able to iterate and improve their performance in some learning activities. The good performance in the continuous assessments suggests that students have indeed been highly motivated to engage in these learning activities and that the graded learning activities with focus on feedback to both students and teachers do have the potential to change students' behaviours and learning patterns. Obtaining conditions where deep learning is maximised and performance in assessments is free of anxiety appears however not to be a straightforward task. It is for example possible that the continuous assessment activities should be given to students in a different format during the course to increase the motivation for engagement in the final exam. It is also possible that a final on-site exam should be completely avoided to minimise the negative effects of anxieties on assessment results. We await the result of focused student interviews to answer these questions.

Ole Eggers Bjælde teaches at all levels at Aarhus University and is developing both teaching and assessment methods for a more innovative, effective and modern practice. Ole has a background as an astrophysicist and is, among other courses, teaching a first-year course within the physics programme from which an example in this paper comes.

Tove Hedegaard Jørgensen works in the area of evolutionary genetics and ecology and has many years of experience in teaching these subjects at several European universities.

Annika Büchert Lindberg teaches assistant professors and PhD students in active learning, course design and innovative assessment methods. She is developing blended learning with a special focus on continuous assessment. Annika has been the project manager for a capacity building project in Vietnam and the International Biology Olympiad in Denmark.

Literature

Andersen, H.L. & Tofteskov, J. (2016). Eksamen og eksamensformer. Betydning og bedømmelse. Samfundslitteratur, Frederiksberg

Andrade, H., & Du, Y. (2007). Student Responses to Criteria Referenced Self-assessment. *Educational Administration & Policy Studies Faculty Scholarship*. 1.

Bassey, M. (1971). The Assessment of Students by Formal Assignments. Wellington: New Zealand University Students Association.

Bearman, M., Dawson, P., Boud, D., Hall, M., Bennett, S., Molloy, E., & Joughin, G.. (2014). "Guide to the Assessment Design Decisions Framework." http://www.assessmentdecisions.org/guide/.

Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall M., & Molloy, E. (2016). Support for assessment practice: developing the Assessment Design Decisions Framework. *Teaching in Higher Education*. http://dx.doi.org/10.1080/13562517.2016.1160217.

Bennett, S., Dawson, P., Bearman, M., Molloy, E. & Boud, D. (2016). How technology shapes assessment design: Findings from a study of university teachers. *British Journal of Educational Technology*, Vol 48, no 2, 2017 pp 672-682.

Bjælde, O. E., Caspersen, M. E., Godsk, M., Hougaard, R. F. & Lindberg, A.E. (2015) Learning design for teacher training and educational development. In *Proceedings of the 2015 Australasian Society for Computers in Learning and Tertiary Education (Ascilite), Curtin University, Perth, Australia.*

Bjælde, O.E. & Najbjerg R. B. (2017). Innovativ formidling af førsteårsstuderende som et design-based research-forløb. Læring og Medier, Årg. 9, Nr. 16.

Black, P. & Wiliam, D. (1998). Assessment and classroom Learning. *Assessment in Education*, Vol 5, no 1, pp. 7-74

Boud, D (2007). Reframing assessment as if learning were important. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education* (pp. 14-25). London: Routledge.

Boud, D. & Falchikov (2007a). Introduction - Assessment for the longer term. In D. Boud & N. Falchikov (Eds.) *Rethinking assessment in higher education* (pp. 3-13). London,: Routledge.

Boud, D. & Falchikov (2007b). Developing assessment for informing judgement. In D. Boud & N. Falchikov (Eds.) *Rethinking assessment in higher education* (pp. 181-197). London: Routledge.

Boud, D. & Falchikov (2007c). Assessment and emotion: the impact of being assessed. In D. Boud & N. Falchikov (Eds.) *Rethinking assessment in higher education* (pp. 181-197). London: Routledge.

Bridges, P., Cooper, A., Evanson, P., Haines, C., Jenkins, D., Scurry, D., Woolf, H. & Yorke, M. (2002). Coursework Marks High, Examination Marks Low: Discuss, *Assessment & Evaluation in Higher Education*, 27:1, 35-48.

Bridges, S. M., Wyatt-Smith, C. M. & Botelho, M. G. (2017). Clinical assessment judgements and 'connoisseurship': Surfacing curriculum-wide standards through transdisciplinary dialogue. In Carless et al (Eds.) *Scaling up assessment for learning in Higher Education*, (pp.67-80) Singapore: Springer,

Brown, G., Bull, J. & Pendlebury, M. (1997). Assessing Student Learning in Higher Education, London: Routledge.

Brown, S. (2005) Assessment for Learning. *Learning and Teaching in Higher Education* (1). pp. 81-89. ISSN 1742-240X.

Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57–66.

Carless, D. (2015). Excellence in University Assessment. London: Routledge.

Christensen, H. B. (2016). Teaching DevOps and Cloud Computing using a Cognitive Apprenticeship and Story-Telling Approach. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education* (ITiCSE '16). ACM, New York, NY, USA, 174-179. DOI: http://dx.doi.org/10.1145/2899415.2899426.

Dawson, P., Bearman, M., Boud, D., Hall, M., Molloy, E., Bennett, S. & Joughin, J. (2013). Assessment might dictate the curriculum, but what dictates assessment?. Teaching and Learning Inquiry, 1, 1, 107– 111.

Dawson, P. & Henderson, M. (2017). How does technology enable scaling up assessment for Learning? In Carless et al (eds.), *Scaling up assessment for learning in Higher Education*, (pp.67-80) Singapore: Springer,

Denny, P., Luxton-Reilly, A., & John Hamer (2008). The PeerWise system of student contributed assessment questions. *Proceedings of the tenth conference on Australasian computing education - Volume 78* (ACE '08), Simon Hamilton and Margaret Hamilton (Eds.), Vol. 78. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 69-74.

Dochy, F; Segers, M. Gijbels, D. and Struyven, K. (2007). Assessment engineering. Breaking down barriers between teaching an learning, and assessment. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education* (pp. 14-25). London: Routledge.

Falchikov, N. & Boud, D. (2007). Assessment and emotion. The impack of being assessed. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education* (pp. 14-25). London: Routledge.

Gibbs, G. & Lucas, L. (1997). Coursework assessment, class size and student performance:1984-94. *Journal of Further and Higher Education*, Vol 21, no 2, pp 183-192, DOI: 10.1080/0309877970210204

Gijbels, D., Van de Watering, G. & Dochy, F. (2005). Integrating assessment tasks in a problem-based learning environment, Assessment and Evaluation in Higher Education, Vol 30, No 1, pp. 73-86.

Glofcheski (2017). Making assessment for learning happen through assessment task design in the law curriculum. In Carless et al (Eds.) *Scaling up assessment for learning in Higher Education*, (pp.67-80) Singapore: Springer.

Glover, C. & Brown, E. (2006). Written Feedback for Students: too much, too detailed or too incomprehensible to be effective?, Bioscience Education, 7:1, 1-16, DOI: 10.3108/beej.2006.07000004.

Godsk, M. (2013). STREAM: a Flexible Model for Transforming Higher Science Education into Blended and Online Learning. Paper presented at E-learn 2013, Las Vegas, United States.

Harvard Magazine (2010). John Harvard's Journal: Bye-bye, Blue Books? http://harvardmagazine.com/2010/07/bye-bye-blue-books (Feb 25, 2017).

Hernandez, R. (2012). Does continuous assessment in higher education support student learning? *Higher Education*, Vol 64, No 4 (October 2012) pp. 489-502, Springer.

Heywood, J. (2000). Assessment in Higher Education: Student Learning, Teaching, Programmes and Institutions. London: Jessica Kingsley.

Marsh, H.W. (2001). Distinguishing between good (useful) and bad workloads on students' evaluation of teaching. *American Educational Research Journal*, 38, 183-212.

McDonald, B. & Boud, D. (2003). The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations. *Assessment in Education*, 10(2), 209–220.

Ministerial Order no. 1062 of 30 June 2016 on University Examinations and Grading (the Examination Order). http://www.au.dk/en/about/organisation/index/5/56/5601theexaminationorder/ (Feb 25, 2017).

Park, C. (2003). In other (people's) words: Plagiarism by university students – literature and lessons. *Assessment & Evaluation in Higher Education, 28(5)*, 471-488. http://dx.doi.org/10.1080/02602930301677.

Race, P. (2014). Making learning happen. A guide for Post-compulsory education. Los Angeles: Sage.

Raaheim, A. (2016). Eksamensrevolusjonen. Oslo: Gyldendal.

Ramsden, P. (2003). Learning to teach in higher education (2.nd ed.) London: Routledge/Falmer.

Richardson, J.T.E. (2015). Coursework versus examinations in end-of-module assessment: a literature review, *Assessment & Evaluation in Higher Education*, Vol 40, No. 3, 439-455.

Schraw, G. (1998). *Instructional Science,* Vol. 26, 113–125.

Shields, S. (2015). 'My work is bleeding': Exploring students' emotional responses to first-year assignment feedback. *Teaching in Higher Education*. DOI: 10.1080/13562517.2015.1052786

Simonite, V. (2003). The Impact of Coursework on Degree Classifications and the Performance of Individual Students, *Assessment & Evaluation in Higher Education*, 28:5, 459-470, DOI: 10.1080/02602930301675.

Trotter, E. (2006). Student perceptions of continuous summative assessment. *Assessment & Evaluation in Higher Education,* Volume 31, 2006 - Issue 5.

Uddannelses-og Forskningsministeriet, 2017. Nye tal på kvalitet i uddannelserne, 02. marts 2017 (http://ufm.dk/aktuelt/pressemeddelelser/2017/nye-tal-pa-kvalitet-i-uddannelserne)

Ulriksen, L. (2014). God undervisning på de videregående uddannelser. Frydenlund, Frederiksberg.

University of Western Australia (2015). University policy on: Assessment, policy no. UP15/5, Approved on 02.12.2015. (http://www.governance.uwa.edu.au/procedures/policies/policies-and-procedures?method=document&id=UP15/5)

Yorke, M., P. Bridges, and H. Woolf (with D. Collymore-Taylor, A. Cooper, V. Fitzpatrick, C. Haines, D. Jenkins, and D. Turner). 2000. "Mark Distributions and Marking Practices in UK Higher Education: Some Challenging Issues." *Active Learning in Higher Education* 1 (1).