REPORTS ON TEACHING AND LEARNING IN HIGHER EDUCATION|
2019:8

Introducing Oral Summative Assessment in the Medical Radiation Physics Program at Stockholm University- Concepts, Challenges and Concerns

Bassler, Niels

Stockholms universitet

Centre for Advancement of University Teaching and Learning

Introducing Oral Summative Assessment in the Medical Radiation Physics Program at Stockholm University

Concepts, Challenges and Concerns

Niels Bassler

Medical Radiation Physics, Dept. of Physics, Stockholm University, Stockholm, Sweden

Department of Oncology and Pathology, Medical Radiation Physics, Karolinska Institutet, Stockholm, Sweden

Dept. of Experimental Clinical Oncology, Aarhus University Hospital, Aarhus, Denmark.

niels.bassler@fysik.su.se

Abstract: This report summarizes a proposed pedagogy project which investigates the use of oral examination for summative assessment at the Department of Physics at Stockholm University (SU). Advantages and disadvantages are discussed, brought in context to existing literature, and an implementation plan is proposed.

The proposed project is based on the two courses FK5031 “Radiation Dosimetry” (7.5 ECTS) and FK8030 “Radiation Protection and Environmental Radiology” (7.5 ECTS) which are provided within the Medical Physics teaching programme at SU. The courses are currently offered in a traditional format consisting of lectures, tutorials and laboratory exercises. Currently, summative assessment is done at the end of the course, in the form of a 5-hour closed-book exam.

The exams typically yield poor outcome, where not seldom half of the students fail. Consequently, a course transformation is indicated. Course transformation encompasses multiple aspects of the course design, however, the main focus of this report is on summative assessment. Alternative assessment formats are briefly discussed as well and related to literature findings. The hypothesis is raised that oral examinations may spur students to deeper understanding, and that it may reduce the workload for the teacher.It is concluded that the oral examination format is a promising option for summative assessment and should be tested in practice.

Nyckelord: oral examination, summative assessment

Rapporter om undervisning och lärande i högre utbildning 2019:8

ISSN 2003-1688

Background

This document purports to elucidate, whether oral examination as summative assessment could be a solution to several issues identified with two courses given within the Medical Radiation Physics teaching programme.

The two courses are FK5031 “Radiation Dosimetry” (7.5 ECTS) and FK8030 “Radiation Protection and Environmental Radiology” (7.5 ECTS), and these are a part of the Medical Physics programme, which is provided by the Medical Radiation Physics group (MSF) at the Dept. of Physics at Stockholm University (SU). These courses are given to 3rd and 4th year students, respectively, and typically ~ 10 students per course to follow these courses every year.

The final exam for both courses consists traditionally of a 5-hour, supervised, closed-book exam. The exam set typically consists of 30% numerical questions and 70 % theoretical questions. Only help allowed are simple calculators (provided by MSF), physics handbook (which students typically bring themselves), and the medical radiation physics course table collection.

The exam for both courses are notorious for poor yields, where not uncommonly half of the students may fail. With the traditional exam format, several issues are observed:

1. When students answer the theoretical questions, it is often unclear what level of understanding they have reached. Some answers are very laconic, or verbatim print of memorized sections of the teaching material. Though, repeating memorized material, even if technically correct, does not demonstrate understanding.

2. The limited amount of helping items allowed at the exam may further encourage memorization rather than focusing on understanding.

3. The numerical calculations are “type problems”, i.e. students have been trained to solve this particular type of exercise, with limited amounts of variation. As a result, students seem to concentrate on memorizing any possible equation relevant to be able

to solve these questions, with little focus on understanding these equations.

> As an example, in the radiation protection course, students are supposed to be able to calculate how radionuclides are retained and excreted in a biological system. In a typical exam questions they are presented with a fairly simple situation such as a nuclear incident where a certain amount of radionuclides are accidentally or intentionally ingested into a human body. This type of question appears at every exam and is fairly constant, only varying secondary parameters, such as changing the nuclide (e.g. from ^{137}Cs to ^{3}H), changing the ingestion time (e.g. intake all at once, or continuous intake over some period of time), changing the compartment of interest(a human, animal, or discharge into a container, pond, bassin…) etc. The underlying methods do not change essentially. Students are supposed to apply the relevant relationships which describe excretion and retention in the compartment, and may eventually calculate the total committed lifetime dose by integrating the total number of decays which happened in the compartment.
>
> Students, who have not understood what they are doing, are typically picking an equation from memory which may (or may not) include the parameters given by the question. A correctly memorized, but unrelated, equation may be stated, numbers are filled in, units may be completely wrong, and the (wrong) result may be left uncommented. Worst case is when the student even fails to comment that the result is orders of magnitudes wrong (i.e. ingesting a trace amount of a nuclide leading to a 1000x lethal dose), since this also shows that the student has no clue on the magnitudes of relevant physical parameters.

Here, I would like to emphasize that the equations in their differential form are very straightforward, and directly convey the physics reasoning behind them, i.e. are the closest tool to demonstrate an understanding of the problem. If a student cannot establish these equations, there is very little hope that the student will reach any profound understanding successively. However, students tend to focus on the mathematical derivation of these differential equations, or maybe simply asserting the memorized integral form (which might have been the solution for a different situation), and then apply this (possibly wrong) form to obtain some result - after all, this is what the question was about?
This problem is already addressed on the course format side, but the hypothesis is raised in

this proposed project, that changing the summative assessment form may further compel students to deeper understanding.

4. Also for the numerical calculations, much exam time is spent on putting in numbers into a (non-programmable) calculator. Often it is necessary to interpolate from tables, which is prone to typing errors. Those errors then propagate throughout the entire problem to be solved, and the student loses precious time when recalculating these figures. The numerical problems to be solved are rather straightforward to solve using spreadsheet or programming scripts. Using simple non-programmable calculators to solve these questions, seem anachronistic today and out of place as it is not intended that the exam should test how fast the students can interpolate manually. The exam should concentrate on probing knowledge and understanding instead.

5. Allowing students to use computers during exam to alleviate the issue mentioned in the previous point is judged problematic, as it is difficult to avoid internet access and thereby collusion among the students. The latter would violate principles about individual assessment rules.

6. Furthermore, the closed-book form encourages cheating, and it is in fact straightforward to cheat in the given setting. Students bring their own lookup tables, which are never checked for any notes which might have been added. Equations and scripts could be printed e.g. soft-drink bottles with small print. Mobile phones can be hidden outside the exam room when students go to the toilet. Even if students sit fairly spaced apart, they could still copy text from each other.
   That said, I have never personally witnessed any clear evidence hereof - however this has happened in the past. Rarely, I observe written answers, which raises the suspicion that the student has had access to material which is not allowed during the exam.

7. The exam format is very time consuming for the teacher, and possibly could be optimized in a more efficiently using another exam format. Preparing a new closed-book exam set typically takes 4 full work days (integrated over all contributors). Answers from the students are typically 150 pages in total, and correcting these takes normally 3-4 full work days. In total, an exam is about 1.5 week of full time work.

Two courses + two re-exams, this amounts one full month per year, or more. Naturally, this leads to recycling of previously used problems which seem to have happened extensively in previous years (even with a complete identical exam set being used with only 1 year difference).

8. In addition to the previous point, many of the old exam questions feel “constructed” and abstract with very little real life motivation. This could be improved. A typical exam question in its traditional format may be:
    - *“What advantages and disadvantages does a silicon diode detector have relative to ionization chambers?”*

Relevance could be increased though, if the above question is changed to:

- *“Ionization chambers are widely used for radiation dosimetry at radiotherapy departments. Alternatively to ionization chambers, silicon based semiconductor diode detectors could also be used. Assume you work at a new radiotherapy department, and the head of the department considers to buy a diode dosimeter. You are given the task to write a short report (~1 page) to your superior, where you explain when (and why) you would use a semiconductor diode dosimeter instead of an ionization chamber, while highlighting any advantages and disadvantages diode dosimeters may have. Your answer should also include any relevant aspects of cavity theory.”*

As mentioned before, the student answers (with few exceptions) typically demonstrate disappointingly low levels of understanding, with a focus on blindly applying methods for certain types of problems, which have been trained during the exercises or by reading previous exams.

The vision for a successfully transformed course is that all students, irrespective of their background and individual learning styles, should

- achieve the intended learning outcomes
- have a profound understanding of the concepts associated with the learning outcomes
- realize that understanding is superior to memorizing concepts.

Successful course transformation will be reached, once students are able to pass the summative assessment demonstrating *good* understanding of the intended learning outcomes.

Course transformation involves redesigning the teaching format on multiple levels. The course transformation has been under evolutionary development (more on this later), since I took over the courses in 2016.

What has been left untouched so far, is the summative assessment for those courses, which will form the main topic of this report. The aim of this report is to investigate which summative assessment forms further could aid the transformation process. Will introducing oral examination at the course programme help to achieve and probe higher levels of understanding? What experience can be found in literature? What disadvantages may there be?

Possible causes

To establish a proper base for appraising the various examination formats, it is worthwhile to scrutinize what may be the reason for the observed issues.

The traditional course format persisted for decades, and may have worked well in the past. However, over the last three decades, a much broader spectrum of students with different cultural background and skills appear at universities. This may be attributed (among other factors) to the Bologna process. To understand this, it is worthwhile to take a closer look at why the Bologna process was conceived.

The Bologna process, which was started by an international agreement in 1999, is a large international higher education reform which has wide impact on how teaching is conducted at Universities in Europe and beyond. The purpose of higher education is stated in the joint declaration as “preparation for the labour market, preparation for life as active citizens in a democratic society” [1] and seeks also to develop and maintain a broad advanced labor base. This increases the influx of students to the higher education institutions.

A key concept of the Bologna process is the mobility of students across national borders. A set of common requirements for all universities which must be fulfilled to ensure intercomparison. By the Bologna declaration, member states agree to harmonize their higher educational system and adhere to an international credit transfer system. The workload and subsequent summative assessment of the curricula are quantified using the ECTS workload and grading system. This added mobility, further widens the cultural spectrum of the students at higher education institutions.

The Bologna agreement emphasizes the definition of learning outcomes, and introduces the Dublin descriptors for aligning the qualification [1]. Specifically, the Dublin descriptors are build on

- knowledge and understanding
- applying knowledge and understanding
- making judgements
- communication skills
- learning skills

This raises numerous questions: When does a student understand a topic? How can the understanding of a topic be assessed? And most importantly: how can I as a teacher help the student to understand?

The last question in particular is challenged due to the Bologna reform. Historically, higher education used to be only available to a small elitist and homogenous fraction of the population. Most information transport from the lecturer to the students happened in the form of one-way lectures. Today, a much broader segment of the population enter the higher education system, with very heterogeneous skill sets. This puts new demands to how higher education is realized, moving the focus from the teacher to the student [2].

The Bologna working group acknowledges this by highlighting student-centered approach and a general shift from teaching to learning. A premise is here set, that successful achieving the learning goals is a consequence of deep-learning which unlocks higher cognitive levels of understanding. Rather than merely assimilate knowledge, the students are expected to be able to relate and apply the pieces of knowledge which belongs to the curriculum. Alignment between the intended learning outcomes and summative assessment is thus indispensable [1,2].

In this particular case with the Radiation Dosimetry course and the Radiation Protection and Environmental Radiology course, the observation is made that the previous examination format encourages memorization more than understanding. The course transformation envisioned here should aim to restore focus on understanding, e.g. by scaffolding [2]. Identifying proper summative assessment format and implementing these cannot be a cure in its own, as transformation must happen on multiple levels. In this report, the summative assessment is carefully considered in a more holistic view, which will be described below.

Key Concepts

Several summative assessment formats have been investigated throughout this project, with emphasis on the oral examination (also sometimes referred to as "*viva voce*" Pearce and Lee, [3]). Other formats investigated are e-Assessment and open-book examinations.

Oral Examination

Attention was given to the concepts of reliability and validity [2-6] of oral examination format. Reliability is defined e.g. in [2] as how the assessment can be trusted, that is, does the assessment faithfully represent the knowledge of the student. Specifically, the summative assessment must be designed in such a way that they reliably probe the level of understanding of the intended learning outcomes [1,2].
Validity concerns with whether the assessment actually is assessing what it is supposed to assess. While outlining the traditional exam forms in the previous section, it was mentioned that an unreasonably much time is spent on error-prone trivial tasks such as interpolating tables using minimalistic calculators. The interpretation is that this weakens the validity of the exam format. The reduced validity also impacts on reliability, by compromising it as mentioned in [4].

Van der Vleuten and Schuwirth [5] takes these definitions of reliability and validity further and quantifies it with a number which can be determined if a significant statistical base is at hand, that is, if the student groups are sufficiently large. This is not possible in this report, since participant numbers are typically ~ 10 or less. Nonetheless, when applying this metric in large classes, both [5] and [6] point out that oral examination tend to have poor reliability scores. Van der Vleuten and Schuwirth [5] emphasizes that this does not necessarily mean that oral examination itself is an inferior examination form, as other factors may have to be taken into account. For instance, [6] mentions that reliability can be improved by involving multiple examiners.

Turner *et al.* [6] conducted oral summative assessments with over 100 respondents. Oral assessment were regarded as a "refreshing alternative to essays", for the students, but also describes that it may only suit certain people. More than half of the students were positive inclined for oral assessment. Those not comfortable with this format, still regarded it as supportive for their learning, as they realize that presentation skills are needed for professional development. This underlines the importance of motivating to the students the reasons for implementing this exam format.

Turner *et al.* [6] also mentions constructive alignment [2], how oral assessment helps to develop knowledge further in ways written assessment cannot. Affective factors are highlighted in the same study, as students may find the oral assignments as "nerve wracking" and "stressful". Possibly, in the case of Turner et al., this is amplified by the selected oral examination format, where peers were present as audience during the examination.

Pearce and Lee [3] worked with oral assessment for marketing courses. Even if not *Natural Science*, this reference is included here, as it summarizes well the pro's and con's for oral assessment substantiated by several references. On the positive side, they list that oral assessments give the student a better opportunity to demonstrate their strengths and provide a better impression of the students knowledge and abilities, than written assessment formats, which is in line what Turner *et al.* reports in [6]. The interactive nature of oral assessments allows examiners to discriminate between superficial and real knowledge by questioning. Related to this observation is that feedback is an integrated part of the oral exam: McCarthy [7] argues that feedback should not only be a part of formative assessment (see Yorke [8]), but also of summative assessment.

Criticism of oral assessments, however, cover mostly psychological aspects and in particular the aforementioned low reliability and validity [3,5,6]. Reasons mentioned for the inferior reliability and validity are e.g. possible bias (oral exams are by nature not anonymous) and poor interexaminer reliability [3]. In the same paper, references are given emphasizing that examiners of oral assessments must be "carefully selected, trained and monitored to avoid allocating marks that have no bearing on competence". Trust to the examiner is important, to avoid defensive or aggressive student behaviour. Preparationary techniques are suggested which may help to reduce the students' stress levels, "such as guidance from the supervisor, clearly defined guidelines, and mock vivas". To increase reliability, Pearce and Lee lists references which suggests to "use several vivas, use several examiners, ask candidates the same questions, use descriptors, employ rubrics and criteria for answers and train examiners". Biggs and Tang also suggests to tape record the oral assessments to avoid possible disputes [2, p.272].

Cheating at oral examinations is virtually impossible, which also indirectly is mentioned by Biggs and Tang, who highlight oral assessment as a way to reduce plagiarism [2, p. 270]

In the case of this pedagogy project, oral examination on its own may be insufficient to assess the learning outcomes. As mentioned earlier, the curriculum of both of the courses requires the students to be able to do numerical calculations, which are difficult to realize during an oral examination. The numerical part, however, could possibly profit from a different exam form which better resembles a real-life situation. Possible solutions from the literature study considered are the open-book exams and e-assessments, which are covered in the next section. Since the theory part of the intended learning outcomes are assessed orally, the remaining numerical part could in principle be assessed by a supplemental open-book exam.

Open-Book Examination and/or E-Assessments

Open-book exams are briefly commented in Biggs and Tang [2 p.228]. Ideally, this open-book exam would be of the "Bring Your Own Device" (BYOD) type, where the students may bring their own laptop (or tablet) with any information they need, having access to spreadsheet programs, programming/scripting environments, etc. Such IT-augmented exams are in line with the idea of blended learning [9], and seem to most naturally recreate a realistic situation where their knowledge should be applied. In fact, due to the widespread adoption of information technology, and its inclusion into teaching in the form of blended learning, it seems anachronistic not to fully endorse these technologies: rather excluding IT from summative assessment, there should be ways of how to assimilate it.

The e-assessment form is highlighted by [10,11]. Indeed e-assessment may be useful e.g. for remote learning, which are common in sparsely populated areas [10], and it is foreseen that remote examination will become more frequent as courses are "digitized", and more courses may be provided as online courses.

Open-book exams eliminates several possible vectors of cheating, but unfortunately, BYOD type open-book exams would enable collusion by wireless communication, which is practically impossible to prevent. This is also valid for e-assessments conducted remotely. Individual assessment are required by SU[1], as they adhere to the Bologna agreement [1]. A rather recent paper (2018) by Güningen et al. [11] goes into some technical detail on current e-assessments and how to avoid cheating. None of the presented techniques where found useful however in such a small setting, as they are either too extensive in terms of technical setup and/or man-power. To prevent collusion, the reference suggests to carefully monitor the students in the examination room, possibly video recording it or by installing surveillance software on the computers.

To start with the latter, installing surveillance software is entirely disregarded for multiple reasons:

- requiring installation of third-party programs on BYOD will most likely violate GDPR regulations. This is an often overlooked fact, and will rule out most supervised e-assessment methods.
- BYOD cannot dictate platform homogeneity. Users may bring computers running on Windows, MacOS or various flavours of the Linux operating system. Very few systems are platform-independent, such as using VNC protocols[2], and still these require much setup work prior to the exams. Furthermore, they increase the risk of failure, in case e.g. network problems.
- Running e-assessments via a network browser is platform independent, but since these are sandboxed in the browser environment, there are no (legal) ways to check for collusion beyond what happens in the page loaded by the browser
- Providing pre-installed computers may on the other hand limit the user to the existing tools at hand. Some users prefer working with Matlab (which requires a live network connection), some with Python, some in R or something completely different. Some prefer MS Word, some are die-hard Linux fans (- which I personally can relate to very

[1] Group exams and collaboration are in principle possible, as long as the individual contributions can be assessed. See section 2.2.2.1 in "Regler för utbildning och examination på grundnivå och avancerad nivå", Stockholm University.
https://www.su.se/polopoly_fs/1.434695.1557829685!/menu/standard/file/Regler%20f%C3%B6r%20utbildning%20och%20examination%20p%C3%A5%20grudniv%C3%A5%20och%20avancerad%20niv%C3%A5.pdf

[2] https://www.hep.phy.cam.ac.uk/vnc_docs/index.html

well: I would feel very encumbered myself should I be forced to work on a Windows/MS Word platform). One may also ask, if this is not a step back from the philosophy of exams should ideally mimic real life situations.

The former solution, video-surveillance, requires video cameras of sufficient resolution and straight line of view to see all computer screens. Most likely multiple video cameras have to be setup, and again this would require unreasonable much work, and still not eliminate ways of cheating.

Alternatively, an open-book summative assessment without BYOD could be considered. Open-book exams stipulates that the focus of the students should be understanding rather than “memorizing and applying equations”, which is the goal of the course transformation envisioned here. This exam format could readily be implemented, since the theory part is covered by oral assessment. Unfortunately, this will probably mean that the total workload would not decrease (a careful estimate is that it would increase).

In summary,

- for technical and possibly legal reasons, e-Assessment is not considered as an option,
- and neither are open-book exams with BYOD.
- Open-book exams without BYOD may be an option, but will require additional studies.

To limit the scope of this report, the emphasis will be put on the design of the oral assessment component in the remaining report.

Design

Apart of the literature study, the design for summative assessment presented here, also relies on experience collected through 3 years of teaching at MSF, and personal experience from examining ~100 students at Aarhus University orally over 3 years (different courses, though).

Context

Since taking over the responsibility of the two aforementioned courses at SU in 2016, I have gradually introduced new elements for an evolutionary course transformation. An evolutionary course transformation is chosen, since the issues need to be carefully understood first, before a transformation can be developed. Furthermore, a radical transformation is time consuming in preparation, requiring additional teaching hours which are not available. Some of the new elements I have introduced to the courses are relevant to explain the ideas behind the chosen summative assessment format, so they will be briefly summarized below.

Alignment with the Intended Learning Outcomes

Both courses were from the beginning well aligned with the intended learning outcomes. The significance of this seemed to be unclear to the student, so I decided to start every course with a few minutes to disclose and briefly discuss each intended learning outcome. I highlight the significance of the level of understanding which is expected, based on the SOLO-taxonomy, and also how this is linked with the final exam. The students are also told, should they during the course be overwhelmed by the amount of information, they can turn to the intended learning outcomes in the course curriculum, to regain orientation. Once the intended learning outcomes have been presented, the course schedule can be motivated to the students as the path to the intended learning outcomes are obtained.

This is based on the idea of constructive alignment, and gives already the first clue to the student that actual understanding the topics is essential, and eventually they must demonstrate their understanding at the exam. As an example to the contrary, I mention an actual case, where a student was able to memorize all slides (or possibly cheated) and provided these memorized text fragments as written answers at the closed-book exam. I explain to the students that reciting memorized text - even if technically correct - does not prove in any way that they understood what they wrote. At the introduction lecture, a small quiz is also given to the students, which is designed to underlines the idea that understanding is superior to memorization.

Lectures

The traditional lecture form is mostly retained, as this forms a stabilising element throughout the transformation process. A few relevant changes are mentioned here, though:

- Course content is “atomized”, i.e. encapsulated in smaller learning bits, when possible, keeping the scaffolding concept in mind. With these atomized content bits, it is very straightforward to implement formative assessment techniques. Lectures are started with a 10-20 minute informal quiz, where the lecturer asks the students *ex tempore* about the most essential parts from the last lecture. This has several beneficial effects:

  - The students become aware of what is important, and become aware if they have not acquired sufficient knowledge about this topic.
  - The teacher gets instantaneous feedback from the students on whether they actually absorbed the teaching material from the previous lectures. I typically spend additional time to repair any uncovered misconceptions.
  - The dialogue between the students and lecturer also breaks the barrier and by my experience also makes the student more attentive to the lecture to follow immediately. The students are more inclined to ask for elaborations.

- It is widely acknowledged (see e.g. Biggs and Tang [2]) that successful understanding is assisted by properly motivating the context of why a certain topic should be learned. Course material has thus been consolidated: in order to emphasize understanding the most fundamental aspects of the course (which often are missed totally judging from previous exams), less relevant parts have been completely removed from the course.

Exercises/Tutorials

Traditionally, exercises tend to end as teacher monologues. Problem solving in the form of exercises and tutorials, however, may serve as a platform for preparing the student for oral assessment, if the student can be activated.

Student presentation of the exercises is motivated to the students as a chance to train their ability to express themselves using professional scientific language. Equally important is that the students gradually will earn confidence “being on stage” in front of an audience, yet under informal circumstances. Feedback is immediate, and the tutorials can be considered as formative assessment, which will help to guide the student to better understand what is expected from him or her [8].

In practice, to motivate the students to go to the backboard, several changes were enforced:

- Students need to have time to absorb and work themselves with the material just learned in a lecture, before they start solving the exercises. Compared to the old course schedule, all exercise sessions are now delayed by one week relative to the lectures, so there is time to process the new material.

- Students are told that the actual results of the exercises are secondary, and they should not worry if they have wrong or no result at this stage. What matters is understanding the methods on how to get to the result. Students receive clear instructions that they should simply begin to do the exercises, and when they get stuck, spend some time in trying to think and formulate why they got stuck, and what piece of knowledge they need to proceed.
- Students are instructed beforehand that they are supposed to present the exercises at the blackboard. Even if the students are not able to complete the exercise at the blackboard, merely setting up the question at the blackboard is already training their communication skills, and will usually trigger deeper discussions.
- Finally, and equally important, I realized that exercise preparations must be scheduled as an out-of-class activity with time reserved for it in the course schedule. The tutorials are scheduled regularly every wednesday afternoon, and now the mornings are reserved for exercise preparations which are unattended by the teacher. Following the idea of student-centered learning, the students are left to their individual preferences to work in groups or alone.

After implementing these points, most students were eventually able to solve most exercises, and almost all students volunteered to present the exercises at the blackboard. Also, keeping the oral examination in mind, I added additional ad hoc questions to also train the students ability to reflect on the learned material, again with the SOLO-taxonomy in mind, e.g. by reciprocal questioning [2 p.167].
The reformation of the tutorial format is thought to be the most vital part which may lead to a successful course transformation. Due to their dialog-intensive and interactive nature of the reformed tutorials one can speculate that it will work favourably with the oral examination format.

A final word on the tutorials can be made on the psychological barrier between the student and “being on stage” at the blackboard. Compared to what I experienced at Aarhus University (AU), Denmark, students at Stockholm University seem less experienced presenting orally. SU has much emphasis on written exams, at AU exams may equally be oral or written. This increases the demand for the teacher to not only establish a safe environment, but also make the student feel comfortable being on stage. Means are employed to reduce the physical and psychological distance to the blackboard by letting the students think of it more like a tool help express ideas, rather than a stage. For instance, one can get the students involved into discussions where they simply need to plot something in order to express what they mean. This causes an increased movement of students to the blackboard, a more lively discussion, breaking down the invisible wall between the blackboard and the audience. When successful, the student “forgets” they are on stage, and speak freely and unencumbered.

Summative Assessment - Oral Examination

For transforming the two courses discussed here, special attention was directed towards oral examination. As pointed out by the previous literature sources, oral exams may be more effective in probing understanding, as the examinator is able to have a dialogue with the student, giving him/her a chance to clarify any inconsistencies.

I will now discuss in more detail, what oral examination format is envisioned.
The oral examination should replace, at least, the theoretical part of the traditional closed book examination used for the two courses. Several degrees of freedom exist how the exact format should look like, but based on the literature studies and in particular previous experience I gained at Aarhus University, a very specific set of oral examination guidelines can be established. This format should be regarded as “safe starting point”. It has been tested already for a few complementary exams for the two courses. The outcome will be addressed later in the implementation section of this document.

§1. Oral examination will be conducted in 25 minutes, with 5 minutes for evaluating the grade. (cf. [3,6])

§2. Three persons will be available in the examination room:

- the student,
- the main examiner (ME), and
- a third person as an associate examiner (AE).

At least ME or AE should be of similar gender as the student, to counter accusations of gender bias.

§3. The AE's foremost role is to witness and ensure a fair process for the student. AE may object to questions if the ME questions beyond the curriculum. The AE may also ask supplementary questions to the student, but the exam should be led by the main examiner. It is the responsibility of both the AE and ME that the exam situation ensures the best possible outcome for the student.

§4. When the student enters the examination room, the student must be introduced to the people present in the room and the examination process. The student must understand that:

- AE is present for their benefit to ensure a fair process
- The examination is about probing what the student *knows*, and not probing what the student does *not know*. (cf. [3])

§5. Peers are excluded for the examination to reduce pressure on the student. According to Turner et al. deeper preparation may be conducted through the anxiety about the public nature or oral examination when peers are present [6 p. 671]. I disagree on that asseveration principally, anxiety should not be applied as a motivating factor for deeper learning.

§6. The intended learning outcomes will form the base for the examination. A list of topics or questions will be handed out to the students at least one week before the oral examination (or at course start - the intended learning outcomes are the exam questions). The questions will be printed on cardboard cards, one for each question. The cards will be present in the room facing down. When the examination starts, the student will randomly pick one card, which will be the topic of examination. Examination starts immediately.
An alternative format is to give the students 30 minutes of preparation time, while the other examination is being conducted. This format may be useful e.g. if mathematical derivations are among the exam questions. (cf [2])

§7. Students are expected to give a 10-15 minutes autonomous presentation to each of these questions at the blackboard. Examiners may interrupt the student any time and probe the understanding and knowledge. The end of the examination may be used to ask questions in a broader sense, to avoid that the student is "playing lottery" by only preparing for a few of the possible exam questions. The expectation is that this will increase both validity and reliability of the oral examination. (cf. [2-6])

§8. The student may bring one piece of paper with notes with them, as this helps to reduce anxiety. By experience, students will in most cases not use it. If the student, however, consistently is not capable of answering without reading from the notes, this may affect the grade accordingly.

§9. Examiners must at all times ensure a safe and friendly atmosphere. This requires that they are aware of their body-language, and ask in a polite and friendly tone. Questioning should be simpler at the beginning, for the student to acclimatize to the situation. (cf. [3]) On the other hand, if serious gaps of knowledge are encountered, this should be emphasized by the examiners so the student will understand the grade afterwards. Still, examiners should not dwell on gaps in knowledge, and move on to other topics. Again, the purpose of the exam is to probe what the student does know. (cf. feedback [7])

§10. Towards the end of the exam, questions may be asked, which the student does not know the answer to, but where the students may use the acquired knowledge to reflect and hypothesize. This is a way to ensure the students have achieved a high level of understanding (cf. to the SOLO taxonomy [2]). The MA may beforehand prepare questions which have not been discussed in class, but where the students may demonstrate their ability to reason using their acquired knowledge. Even if the student might not reach the correct answer, their ability to reason demonstrates scientific maturity, which must be awarded accordingly.

This examination format is also standard at the Dept. of Physics and Astronomy at Aarhus University in Denmark, even if less specific. One difference is that the AE may be an internal or external person, called a “sensor” (in danish: “censor”). External sensors are chosen from a national list of people entitled to do these examinations, and regulations foresee that one third of all examinations are done with external examiners to ensure homogeneity in grades across the higher educational system[3].

---

[3]See http://phys.medarbejdere.au.dk/fileadmin/site_files/censorformandskab/introduction17.pdf (in English)

Possible future extensions to the here proposed examination format would be to employ the aforementioned rubrics, to clarify even further what level of understanding is required. This is disregarded here, since the intended learning outcome formulation already clearly expresses what is required. That said, over time, such rubrics could be developed, but for now this is regarded as a secondary feature.

Finally, I witnessed some curiosities practiced at oral examinations abroad, which I leave here for the record:

- If the examination happens sitting around a table, the ME should face the student. The AE may sit perpendicular to the student and ME, to emphasize AEs independence of the ME.
- The ME should face the window in the room, so the student facing the ME is not glaring against the light.
- In Denmark, oral examinations are traditionally conducted with a green table cloth made of felt. The green colour is a symbol of hope and is supposed to help the student to relax.

Implementation

At the Medical Radiation Physics group, the course formats have been fairly static. The curriculum and exam forms have undergone few changes, exams from 20 years ago still look fairly similar to those given today. Understandably, both fellow teachers and students may react conservative to the introduction of drastic events such as changing the summative assessment for a course which has demonstrated its worth for many years.

As a starting point, the curriculum for both courses does not preclude oral examination for summative assessment, both exam forms are explicitly allowed. Thus, no change of the course curriculum is needed.

See also http://phys.medarbejdere.au.dk/formandskabet-for-censorkorpset-i-fysik/ for detailed reports about the system (in Danish).

All course lecturers and students should be informed of this change, as this will be the first graded oral examination happening on the medical physics programme. Prior implementation, feedback will be collected from the head of studies, and programme responsible.

Concerns

From a lecturer side, points of concern could be raised, i.e that oral examination:

- may reduce the assessment standards, e.g. by not probing important aspects of the intended learning outcomes,
- may introduce extra work (e.g. since an associate examiner must be present during all exams),
- may provide a breeding ground for more student complaints.

Students, on the other hand,

- may find it concerning to abandon the well-known fixed-format examinations,
- may be uncertain on how to prepare,
- may worry what to expect from this examination format,
- may fear that the examiners cannot be trusted,
- may fear that they will underperform since they are not trained for the oral examination format.

Prior implementation, additional student interviews may be appropriate to elucidate whether additional issues may have been missed.

Workload Analysis

As mentioned earlier, one of the advantages is the anticipated reduced workload of oral examinations. To stipulate this, the following time budget was made (based on either the Dosimetry or Radiation Protection Course, exam workload is fairly equal).

Task	Written examination 10 students [man hours]	Oral examination 10 students [man hours]
Prepare written examination questions (incl. solutions)	20 additional lecturer 1: +2 additional lecturer 2: +1	1 *)
Booking of rooms	0.5	0.5
Arrange exam watch, print exam sets, handouts,...	1.0	0
Examination	5	4.2 by AE: + 4.2
Correction of exams + Grading	15 additional lecturers: +4	0.8 by AE: + 0.8
SUM	**49.5**	**16.0**

*) The questions may recycled every year. Preparation time is not counted here, as this is equal to the preparation for the lectures and tutorials.

Both solutions scale with the number of students, but it is important to stress that the examination workload scales slower with the oral examination, as these only take ~ 0.5 h (= 1 man hour) per student.
Correcting a written exam set (typically 15 pages) may take easily 1-2 hours per exam set. Here 1.5 hours are assumed, and another 40 minutes by two or three additional lecturers correcting their questions, i.e. approximately 2 man hours per student.

What is missing in the oral examination list is the assessment of the numerical part for the examination. Pointwise, these are ~ 30 % of the entire exam set.

Standard teaching workload calculates with 40 hours for the main examiner per year, which must also cover one re-examination. This seems insufficient for the standard written exam, since the actually spent time here is closer to the presented calculation, i.e. 2 x 41.5 = 83 hours of work for the main examiner (not including the contributions from the co-examiners). Compared to the oral examination, the ME, spends merely 6.5 hours per exam, i.e. 13 hours in total assuming one re-examination. The remaining time additional 27 hours which could be spent on assessing the numerical part, or as additional preparation for the oral examination.

Training

Colleagues unfamiliar with this exam format may receive training by sitting in such exams, or by taking the role of the associated examiner. Possibly, this examination form is not suited for every colleague, however it may be important first to experience this examination format before reaching a decision.

Preliminary Test Results

Several actions for reducing these concerns have already been implemented. When the idea of oral examinations was initially presented, colleagues reacted will reasonable skepticism. To resolve this, colleagues participated in small test cases with students who had failed their exams with an “Fx” and needed a complement. The colleagues acted as associate examiner while I conducted the examination. So far, all colleagues reacted positive to it, and they agreed it could be worthwhile to try out this examination form for the courses.

The students I have examined this way also responded positively afterwards. Anecdotally,

- one student who at written exams seem to be unable to express himself by written text, demonstrated surprisingly good understanding during oral examination - to a level I was not able to uncover by his answers from the written exam.
- Another student failed the oral exam. The verdict was readily accepted by the, since questioning clearly exposed the student had not acquired sufficient knowledge and understanding.

Of five students examined at MSF, no negative experiences were made, on the contrary.
I have already described how the students can be trained for the new situation by formative assessment during the exercises and tutorials. One idea could be to use this examination form with the consent of the students: If they do not want oral examination, then they can anonymously vote no. The fact, that up to 50% of the students tend to fail the written closed-book exam may encourage students to try this examination form.

Conclusion

Oral examination as summative assessment appear to be a promising solution to transform the two courses investigated. Among the advantages are better probing of knowledge and understanding, less time consumption for the lecturer, immediate feedback to the student, elimination of possible vectors for fraud and collusion.

Oral examination is compared to open-book examination as an alternative format, and possibly both could be used in conjunction with each other, if resources allow this.

A detailed design is presented on how oral examination can be carried out in practice. Implementation of oral examination cannot happen on its own, but requires also adoption to the entire course format, in particular when students are not used to this exam form.

The design presented in this report is generic, and may hopefully inspire colleagues with similar courses to consider this examination form.

References

1. Ministry of Science, Technology and Innovation. A framework for qualifications of the European higher education area. Copenhagen: Bologna Working Group on Qualifications Frameworks. 2005 Feb. http://ecahe.eu/w/index.php/Framework_for_Qualifications_of_the_European_Higher_Education_Area

2. Biggs JB and Tang C. Teaching for quality learning at university. McGraw-Hill Education (UK); 4th edition, 2011.

3. Pearce G, Lee G. Viva voce (oral examination) as an assessment method: Insights from marketing students. Journal of Marketing Education. 2009 Aug;31(2):120-30. https://doi.org/10.1177/0273475309334050

4. Knight PT. Summative assessment in higher education: practices in disarray. Studies in higher Education. 2002 Aug 1;27(3):275-86.

https://doi.org/10.1080/03075070220000662

5. Van Der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. Medical education. 2005 Mar 1;39(3):309-17. https://doi.org/10.1111/j.1365-2929.2005.02094.x

6. Turner K, Roberts L, Heal C, Wright L. Oral presentation as a form of summative assessment in a master's level PGCE module: the student perspective. Assessment & Evaluation in Higher Education. 2013 Sep 1;38(6):662-73. https://doi.org/10.1080/02602938.2012.680016

7. McCarthy J. Evaluating written, audio and video feedback in higher education summative assessment tasks. Issues in Educational Research. 2015;25(2):153. (no DOI available) https://search.informit.com.au/documentSummary;dn=376762809256535;res=IELAPA

8. Yorke M. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. Higher education. 2003 Jun 1;45(4):477-501. https://doi.org/10.1023/A:1023967026413

9. Garrison DR, Kanuka H. Blended learning: Uncovering its transformative potential in higher education. The internet and higher education. 2004 Apr 1;7(2):95-105. https://dx.doi.org/10.1016/j.iheduc.2004.02.001

10. James R. Tertiary student attitudes to invigilated, online summative examinations. International Journal of Educational Technology in Higher Education. 2016 Dec;13(1):19. https://doi.org/10.1186/s41239-016-0015-0

11. Von Gruenigen D, e Souza FB, Pradarelli B, Magid A, Cieliebak M. Best practices in e-assessments with a special focus on cheating prevention. In 2018 IEEE Global Engineering Education Conference (EDUCON) 2018 Apr 17 (pp. 893-899). IEEE. https://doi.org/10.1109/EDUCON.2018.8363325

Stockholm University 106 91 Stockholm, Sweden
Tel +46 (0)8-16 20 00 www.su.se info@su.se

Stockholms
universitet