ORAL GROUP EXAMINATION METHOD TO EVALUATE COLLABORATIVE AND INDIVIDUAL LEARNING

Karin Stenderup, Sanne Sandberg Overby

Institute of Biotechnology and Chemistry, Faculty of Technical Sciences, Aarhus University, Hangoevej 2, 8200 Aarhus N, Denmark

ABSTRACT

The engineering education at Aarhus University, Denmark, includes the course Statistics and Design of Experiments (DoE) that encompasses both collaborative and individual learning activities. The choice of examination method is important, as students adapt their learning behavior accordingly. Group examinations align well with collaborative learning and individual examinations are successful in assessing individual learning and detecting free riders. To avoid organizing and exposing the students to two examinations and thus imposing additional undesirable costs, we aimed to develop a single highly structured oral group examination method that addresses both collaborative and individual learning in an organized fashion without increasing demands on academic staff. The oral group examination method described in this study is a three-in-one exam where all group members in a project group are present at all times. First, the students' collaborative skills were addressed with focus on knowledge application and analysis. Then their individual skills were addressed with focus on basic knowledge understanding. Finally, students were given the opportunity to evaluate their own knowledge and create new knowledge, which includes the pinnacle of Bloom's taxonomy pyramid. The examination method was tried out with four classes of engineering students (142 in total): two Chemical engineering and two Biotechnological engineering classes in their second and third year. Afterwards, students reflected on their perception of the exam in a survey. In summary, the examination method embraced assessment of both collaborative and individual learning and provided time for in-depth discussions with all group members, in the project group, on a high taxonomic level. We encourage other educators to explore this examination method. The present study includes a "ready-to-implement" protocol and a "ready-to-use" Student Scoring Sheet to keep track of the contribution of each student.

KEYWORDS

Second-Year Undergraduate, Chemical Engineering, Biotechnology Engineering, Testing/Assessment, Student-Centered Learning, Standards: 11

INTRODUCTION

Statistics and Design of Experiments (DoE) is a mandatory course for Chemical and Biotechnological Engineering students at Institute of Biotechnology and Chemistry, Aarhus University, Denmark. The course encompasses both collaborative and individual learning activities. During the collaborative activities, the students are divided into groups, where they actively engage in a problem-based-learning (PBL) activity based on a selected project within DoE. In addition, the students attend traditional lectures and work individually or in groups with exercises concerning theory on statistical methods. Prior to the exam, the students hand in a project report describing their results. Until recently, in order to asses and certify the students' acquired course skills, the course was concluded by conducting a single final individual oral exam.

From 2006-2013, group exams were discontinued in all educational institutions in Denmark. This decision was made due to concern that grade distinctions among students were too blurred in group exams. That is, low-performing students were awarded grades that were too high, while high-performing students were awarded grades that were to low (Krogh & Aarup Jensen, 2013). Thus, individual exams replaced all group exams during this period. In an attempt to assess collaborative learning in courses containing project-organized PBL activities, some individual oral exams would include a group presentation prior to the individual examination of each group member, whereas others would rely solely on an individual examination. The many years without group exams left students and educators, especially the younger ones, unfamiliar with the group examination format. Prior to the reinstatement of group exams, studies among students revealed that students generally did not feel confident about the validity of a group exam (Krogh & Aarup Jensen, 2013). The students worried whether the examiners were able to give reliable and valid individual grades when other group members were present. They also worried that some students would dominate the exam. After years of individual exams the reinstatement of group exams has led to an increased focus on how the students are examined and what they are examined in. Thus, in order to be successfully reinstated, group exams require a renewal process, where group exam procedures are explicated and exemplified more clearly to both students and educators.

The objective of the present study was to develop such an oral group examination method. The examination method should be highly structured and utilize the benefits of the former oral group exam and the current oral individual exam. It should embrace both a group and an individual assessment aspect and ensure the following: 1) time for in-depth group discussions to evaluate the students' ability to collaborate, 2) the ability to assess each student individually, 3) the possibility of detecting free riders, 4) time for the students to reflect on and evaluate their acquired knowledge and create new knowledge, and 5) the possibility of addressing all knowledge levels in Bloom's taxonomy pyramid in a structured manner (Bloom, 1956, 2001). The idea is that more students will demonstrate skills at the pinnacle levels of Bloom's taxonomy when other group members are present to fill in underlying and indispensable knowledge gaps. Finally, the examination method should be conducted without increasing demands on academic staff.

In our highly structured oral group examination method, the total examination time of all students in the group is compiled into one exam, which is then divided into three parts (Figure 1). Part one is a group examination. Part two explores the actual level of understanding of each individual student. Part three is a group reflection, allowing time for the students to evaluate their own knowledge and possibly create new knowledge. Our hypothesis is that this highly

structured oral group examination method will lead to sound and valid assessment of the students' acquired course skills at several levels of understanding.

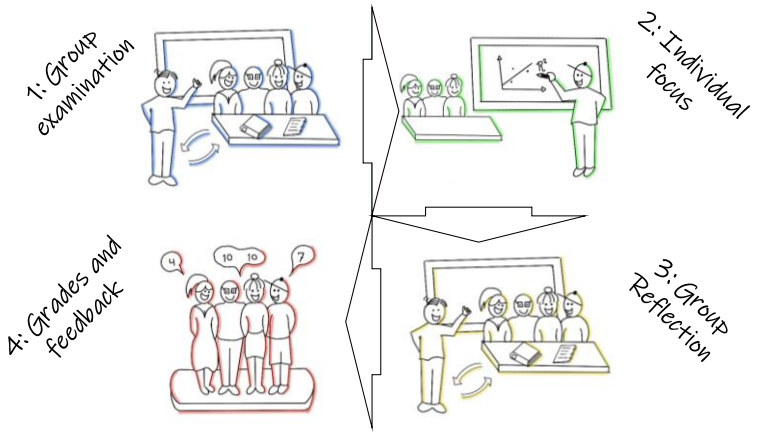

Figure 1. Graphical presentation of the highly structured oral group examination method. 1: shows a group examination where the examiner addresses all students in a group. 2: shows focus on one student to explore individual level of understanding. 3: shows a group reflection where the examiner addresses all students in a group. 4: shows students receiving grades and feedback after the examination

The highly structured oral group examination method was tried out with four classes of engineering students. Subsequently, the students were asked to reflect on their perception of the examination method in a survey. On average, the students obtained grades above a 7 (corresponding to a C) and perceived a maintained ability to perform individually, a valuable experience in being examined both collaboratively and individually in the presence of their group members, and felt that the exam was fair. There was no additional time spent per student, and the examination method provided extended time to address the learning outcomes more thoroughly and, thus, to assess the students in a more satisfactory manner.

In this paper, we present our highly structured oral group examination method, which aligns well with courses comprising both collaborative and individual learning. It also works well with courses with limited time per student for the final examination. In contrast to the traditional individual oral examination, where time is short and the questions are interrogation-like, the authors experienced more time for dialog and thorough group discussions in the oral group examination method, as well as the possibility of addressing higher taxonomic levels of knowledge. All of this is achieved in a highly structured fashion.

Based on the present study, we encourage other educators to explore the highly structured oral group examination method. To assist implementation, this paper includes a "ready-to-implement" protocol (Table 1) together with a "ready-to-use" Student Scoring Sheet (Supplementary material) to ease the task of keeping track of each individual student's contribution during the examination.

CONTEXT

Statistics and Design of Experiments (DoE) is a 5-ECTS (European Credit Transfer System) course, which makes up one sixth of a typical semester. It comprises 14 four-hour teaching

modules, with one module scheduled each week. The first half of the modules are carried out by giving traditional lectures and assisting the students in solving problems, either individually or in groups. These modules focus on theory on statistical methods such as parametric hypothesis testing including analysis of variance (ANOVA) and linear regression, non-parametric hypothesis testing, and how to use control charts to assess whether a process is in a state of statistical control. The last half of the modules are carried out by supervising a PBL activity, where the students are divided into groups and work collaboratively on a selected project within DoE. These modules focus on factorial design and response surface methods (RSM). To complete the project, the groups write and submit a mandatory report of around 10-15 pages prior to the exam. The course learning outcomes address the higher taxonomic levels of knowledge and specify that the students should be able to apply their acquired knowledge, analyze their data, evaluate their results and suggest experiments that would improve their results, it they were to perform the project again.

Four classes of engineering students comprising second- and third-year Chemical and Biotechnological Engineering students (n = 142) attended the course in the spring semester of 2018: two second-year classes, Chemistry [n = 34] and Biotechnology [n = 35], and two third-year classes, Chemistry [n = 31] and Biotechnology [n = 42]. Students were assigned into groups consisting of three to six students: three groups of three students, three groups of four students, 17 groups of five students and six groups of six students. In the second-year classes, students were randomly assigned, and in the third-year classes, students were self-assigned. The learning outcomes were the same regardless of how many students joined a group, whether the students were second- or third-year, or whether the students were Chemical or Biotechnological Engineering students. During the PBL activity, the purpose was to create a final equation describing the variation in a chosen process outcome and furthermore understand the sequential line of events that lead to this equation. The submitted report was a prerequisite for taking the exam, but did not count in the final grade.

At the beginning of the course, the students were informed that the final exam would be an oral group exam, where they would be evaluated both collaboratively and individually, and that grades would be given individually. Course settings allocated a total of twenty minutes per student for the final examination. Accordingly, the highly structured oral group examination method was allocated the total time of the number of students times twenty minutes. The six groups of six students were randomly split into two groups of three students, prior to the exam, in order to reduce the total examination time and to be able to keep track of all students. Thus, the shortest exam was one hour (three students) and the longest exam was one hour and 40 minutes (five students). To illustrate the examination method, a small-scale trial exam was conducted together with the students in class prior to the final exam. Two educators were present for this illustration and volunteer students pretended to take the exam. After the trial exam, the students were given the opportunity to ask questions about the exam.

EXAMINATION METHOD

The highly structured oral group examination method consisted of three parts, in addition to the part where grades and feedback were given (Figure 1 and Table 1). Two academic staff members were present during the exam; one acting as an examiner and the other as a co-examiner. Part one (six minutes per student) was a group exam based on the project report submitted by the students. Here, the students discussed how their acquired knowledge was applied in their project, what choices were made, which obstacles were encountered, and how their data were analyzed. To make sure that all students contributed evenly to the group discussion, the students were instructed to divide the allocated examination time among them

and were encouraged to raise their hand to indicate the possession of a possible answer. The co-examiner, who kept track of time, guided the discussion by asking supplement questions as needed. In this part of the exam, the examiners developed a first impression of each individual students' level of understanding. Following part one, all students left the room while the examiners evaluated the students' midway performances and decided on a strategy to challenge or affirm each individual student's level of knowledge.

After the evaluation intermission all students re-entered the room. Part two (six minutes per student) of the exam then explored the actual individual level of understanding of each student. Each student randomly selected a question for examination from a pool of questions covering the theory on statistical methods taught in the course. In turn, each student was examined. Upon request from the student or the examiners, the student was able to receive help from the group members. Group members were then allowed to rephrase the questions or provide useful hints on how to answer the questions, but they were not allowed to give actual answers. In order to attain positive grading, each individual student still had to answer their own examination question. Thus, this part of the exam covered individual assessment and the detection of potential free riders, which is a benefit of the traditional individual examination. Moreover, group members were present to reduce anxiety, which is a benefit of the traditional group examination.

Part three (three minutes per student) was conducted in the same way as part one, but with focus on reflection, which allowed time for the students to evaluate their own knowledge and possibly create new knowledge. To keep record of the students' responses during the exam, a Student Scoring Sheet (Supplementary material) was employed. If the students raised their hand to indicate the possession of a possible answer or if they rephrased a question or provided a hint to a fellow student, this was recorded in their favor. Following part three, once again, all students left the room and the examiners carried out a follow-up/final evaluation. Finally, part four consisted of grading and feedback, which was given either individually or to the whole group. Altogether, the intermediate evaluation, the final evaluation, grading and feedback were allocated a total of five minutes per student.

All students were graded individually and grades were based solely on their performance during the examination. The Danish "7-point grading scale" system was used, where the grades 12, 10, 7, 4, 02, 00, -03 correspond to the following grades on the ECTS scale: A, B, C, D, E, Fx, and F, respectively. Immediately after grading, the students were asked to participate in a survey, designed by the authors, where they were asked to reflect on their perception of the examination method by filling out a questionnaire. The questionnaire comprised five questions (Figure 2, Q1-5) and the possibility of adding additional comments. The questions were presented as statements, where the students rated their perception anywhere on the scale from "I disagree" (-100%) through "I neither disagree nor agree" (0%) to "I agree" (100%). The statements "I had ample opportunity to say what I wanted" (Q1) and "It felt comfortable receiving individual questions in the presence of the other group members" (Q3) were used to evaluate whether the students experienced a sufficient ability to perform individually. The statement "There was support in receiving and answering questions as a group" (Q2) was used to evaluate whether the students felt supported in being examined through dialog and discussion among group members and examiners. The statements "The examination form was fair" (Q4) and "I prefer this group examination method in comparison with the one I am used to, where the first part is a collective group presentation of the project followed by an individual examination" (Q5) were used to evaluate whether the students experienced alignment between learning outcomes, activities and assessment. It is worth noting that all second- and third-year Chemical and Biotechnological Engineering students had

completed two to three project exams prior to their exam in Statistics and Design of Experiments (DoE) where the format comprised presenting their project collectively as a group prior to a traditional individual oral exam.

Table 1. A "ready-to-implement" protocol for the highly structured oral group examination method for groups of four students

Part	Examination goals	Procedure
1: Group examination 24 minutes (6 minutes/student)	Collaborative skills. (Based on PBL activity)	All students are expected to participate actively in the discussion on how acquired knowledge is applied and used to analyze obtained data.
Evaluation intermission[a]		
2: Individual focus 24 minutes (6 minutes/student)	Individual skills. (Based on theory)	Each student randomly chooses a question and is responsible for answering and demonstrating individual level of understanding. Group members are allowed to rephrase questions or provide useful hints.
3: Group reflection 12 minutes (3 minutes/student)	Collaborative skills. (Based on reflective questions)	All students are expected to participate actively in the discussion and evaluate acquired knowledge and create new knowledge.
Evaluation[b]		
4: Grades and feedback[c]		Grades are given individually to each student with minor feedback or to the whole group with more time for feedback.

Prior to examination, names of all students are recorded on the Student Scoring Sheet (Supplementary material) to keep record of each individual student's responses. [a,b]All students are asked to leave the room and the examiners make a first[a] and a follow-up/final evaluation[b]. [c]Part 4: Grades and feedback and the two evaluations share the remaining examination time: 20 minutes (5 minutes/student). Students are given the opportunity to make commentary notes during the exam. PBL = Problem-based learning.

Statistics

One-sample t-tests were used to test whether the grades were above 7 (C), if the grade distribution followed the normal distribution, and whether the students agreed more to a question in the questionnaire than "neither disagree/nor agree". One-way ANOVA tests were used to compare grades in the four classes (second- and third-year, Chemical and Biotechnological Engineering students, and groups of three, four, or five students). Two-way ANOVA tests were used to take into account whether the students were second- or third-year students and studying Chemistry or Biotechnology. Linear regression analyses were performed to test for correlation between survey responses and grades. $p < 0.05$ was considered significant.

RESULTS

The 142 students who participated and completed the highly structured oral group examination method were graded 7.8 ± 3.1 (mean ± std. dev.) in average, which was significantly higher

than the overall average of 7 (C, p = 0.003). The grades awarded to the four classes were distributed as follows: 4% ± 2% attained grade 2 (E), 24% ± 8% attained grade 4 (D), 28% ± 5% attained grade 7 (C), 25% ± 7% attained grade 10 (B) and 18% ± 8% attained grade 12 (A). No students failed the course (grades 00 (Fx) and -03 (F)). No significant differences were seen in the grades between the four individual classes, between second- and third-year students, nor between students studying either Chemistry or Biotechnology. Although grades are awarded solely based on the student's fulfillment of the learning outcomes, it is expected, for a large population of students, that the grade distribution will follow a normal distribution. In the normal distribution, 10%, 25%, 30%, 25% and 10% attain the grades 12, 10, 7, 4 and 02, respectively. The failing grades 00 (Fx) and -03 (F) are not included in the expectations for a normal distribution. For the four classes participating in the oral group examination method, there were no significant deviations from the expected grade distribution, except that significantly fewer students were given the grade 02 (p = 0.011). In addition, the average grades obtained in the examined groups of three, four, or five students were not significantly different.

Immediately after grading, the students were asked to participate in a survey where they were asked to reflect on their perception of the oral group examination method. All survey responses are summarized in Figure 2. The students significantly agreed on: having ample opportunity to say what they wanted (Q1, 33%, p < 0.000), experiencing support in receiving and answering questions as a group (Q2, 44%, p < 0.000), feeling comfortable receiving individual questions in the presence of the other group members (Q3, 37%, p < 0.000), and experiencing that the exam was fair (Q4, 29%, p < 0.000). When comparing the group examination method with the examination method used in previous semesters in other courses, where students give a collective group presentation of their submitted project report prior to a traditional individual oral exam, the students did not significantly agree or disagree on which examination method they preferred (Q5, 10%, p = 0,061).

The students' grades and their answers to Q1, Q2, and Q4 correlated significantly (Q1: p = 0.001, Q2: p = 0.024, and Q4: p = 0.002). Thus, students receiving higher grades were more likely to agree on having ample opportunity to say what they wanted, experiencing support in receiving and answering questions as a group, and experiencing that the exam was fair (data not shown). Interestingly, second-year students were more likely to agree on having ample opportunity to say what they wanted (Q1, p = 0.031) and on experiencing that the exam was fair (Q4, p = 0.001) in comparison with third-year students (data not shown). In addition, Biotechnological Engineering students were more likely to agree on experiencing that the exam was fair (Q4, p = 0.005) in comparison with Chemical Engineering students (data not shown).

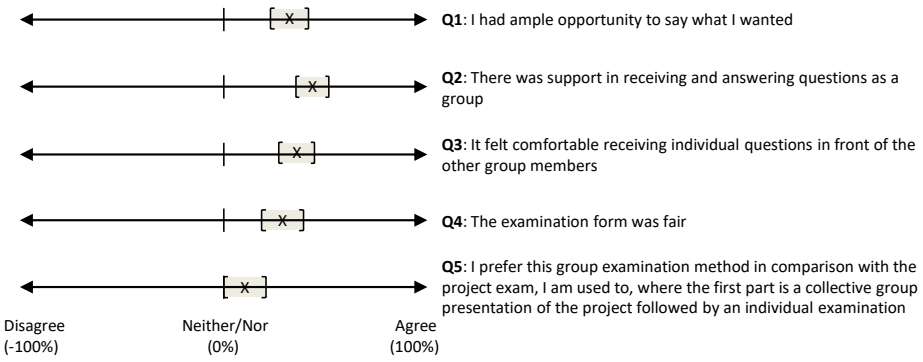

Figure 2. Survey questions and summary. The students were asked to reflect on the group examination method by filling out a questionnaire handed out immediately after grading. All

brackets display the mean response values (X) together with a 95% confidence interval. **Q1:** mean = 33%, [24%;42%]; **Q2**: mean = 44%, [36%;52%]; **Q3:** mean = 37%, [28%;45%]; **Q4**: mean = 29%, [19%;39%]; **Q5:** mean = 10%, [0%;21%]). Data were based on completed questionnaires from all 142 participating students.

In the questionnaire, the students provided several additional comments regarding the group examination method. Of the 142 students: 20 explicitly commented that raising their hand to emphasize the possession of a possible answer worked well, 14 explicitly commented that they liked that other group members could assist them in answering a question, nine explicitly stated that this oral group examination method only worked well, when the dynamics of the group were good; Eight explicitly suggested that the individually chosen questions for the second part of the exam, should not be revealed before time of addressing. This, in order to avoid that group members focus more on their own question and, consequently, pay less attention to the question being addressed by the group member being examined; six explicitly stated that more than five students in a group exam, would be too many.

DISCUSSION

Pedagogical research strongly recommends that course objectives and activities align with the method of examination, as students adapt their learning behavior accordingly (Bretz, 2012; Brown & Glasner, 1999; Momsen et al., 2013; Pienta, 2011; Rhodes, 2010). The objective of the present study was to develop a highly structured oral group examination method that aligned with the current course comprising both collaborative and individual learning. The examination method should meet the demands for clearer explication of the rules for carrying out a group exam and ensure exemplification of the exam format to ensure transparency and confidence among students and examiners. At the same time, the examination method should result in reliable and valid individual grading, without causing additional time-use per student.

Other studies have reported on the combination of group and individual exams, but mostly in a written format where students answer free response problem-solving questions, essay questions or multiple-choice quizzes. In fact, the written two-stage examination method has been widely used for final assessments. In this type of exam, one stage involves an individual part and a second stage involves a collaborative part. For instance, students may hand in an individual response prior to addressing identical or similar questions collaboratively as a group (Levine et al., 2018; Levy, Svoronos, & Klinger, 2018; Lindsley, Morton, Pippitt, Lamb, & Colbert-Getz, 2016; Mahoney & Harris-Reeves, 2017; Rieger & Heiner, 2014; Rivaz, Momennasab, & Shokrollahi, 2015; Vázquez-García, 2018; Wieman, Rieger, & Heiner, 2014). Others have used written group exams as midterm assessments to supplement a final individual exam (Bay & Pacharn, 2017), have offered voluntary group retakes of exams at a later time point following an individual exam (LaBossiere, Dell, Sunjic, & Wantuch, 2016), or have allowed students to work together on a written group exam prior to taking a separate individual exam (Siegel, Roberts, Freyermuth, Witzig, & Izci, 2015). Some studies report a positive effect on the overall average score as a consequence of including a collaborative aspect in the exam (Hanna, Roberts, & Hurley, 2016; Levine et al., 2018; Rivaz et al., 2015; Vázquez-García, 2018), whereas others fail to find an overall effect (Mahoney & Harris-Reeves, 2017), or even find a negative effect (Molsbee, 2013). Interestingly, studies dividing students into high and low performers or dividing questions into basic/concrete and complex/abstract categories report that low performers' grades increase more than high performers' and that answers to complex questions improve more than answers to basic questions when a collaborative aspect is included in the exams (Bay & Pacharn, 2017; Levy

et al., 2018; Mahoney & Harris-Reeves, 2017; Siegel et al., 2015). It is important to note, however, that even high performers are reported to improve from collaboration (Jang, Lasry, Miller, & Mazur, 2017). Moreover, it has been suggested, that group exams improve learning, improve knowledge retention and reduce exam anxiety. Downsides, however, have also been reported and include: 1) retrieval disruption, where students forget due to interruptions, 2) production blocking, where students wait their turn and therefore halt ideas, 3) shared forgetting, where all students in a group overlook the same knowledge, or 4) spreading of misinformation (LoGiudice, Pachai, & Kim, 2015). In addition, some students become stressed when working in the presence of other students (LaBossiere et al., 2016).

In the present study, the students obtained average grades above 7 (C), and no students failed the course. Implementing the highly structured oral group examination method does not seem to have a negative impact on the final grades. It is worth noting, that the group examination part in the highly structured oral group examination method accounts for more than half (6 minutes of collaborative testing plus 3 minutes of collaborative reflection versus 6 minutes of individual testing per student) of the final grade, whereas other studies report a group examination part accounting for only 5% (Levine et al., 2018), 10% (LaBossiere et al., 2016) or 33% (Rieger & Heiner, 2014). In addition, average grades were not significantly different when compared with grades obtained at the individual oral exam used prior to changing to the present group examination method (n = 48 students).

It is unusual, however, that no students failed the course, although, we generally experience a low percentage of students failing this course. We believe that the mandatory report, which is a prerequisite for taking the exam, is of crucial importance. This report sums up the acquired knowledge from the collaborative PBL activity, which covers at least half of the course. If students do not participate actively in the PBL activity, it is difficult to be part of a group report and consequently join the exam. On the other hand, if students participate actively, they spend at least half of the course in collaboration with other students collectively acquiring course knowledge. Moreover, weekly group supervisions most likely improve learning skills through the possibility of discussing difficult learning outcomes with both the group and the teacher.

After the exam, the students were asked to participate in a survey to reflect on their perception of the highly structured oral group examination method. All students responded, rendering great credibility and validity to the interpretation of the response data. Firstly, the students replied that they “experienced ample opportunity to say what they wanted” (Figure 2, Q1), and “felt comfortable receiving individual questions in the presence of the other group members” (Figure 2, Q3), which suggest that being examined as a group did not hinder the students’ experience of being able to perform individually. Secondly, the students “experienced support in receiving and answering questions as a group” (Figure 2, Q2). Undoubtedly, this feeling strongly depends on the existence of well-functioning group dynamics, which was also noted explicitly by nine students in the survey. Interestingly, the authors experienced that, in some groups, resourceful students were able to elevate the performance of the entire group by rephrasing questions or providing helpful hints, but not actual answers during the exam, a phenomenon most likely correlated with favorable group dynamics. Finally, the students “felt the exam was fair” (Figure 2, Q4). This was especially true for students who attained higher grades. That the exam felt fair suggests that the students felt alignment between learning outcomes, activities and assessment. This may correlate with the fact that the students "neither favored nor disfavored” the oral group examination method in comparison with the project examination method, where students give a collective group presentation prior to a traditional individual oral exam” (Figure 2, Q5), as the project examination method also displays both a collaborative and an individual examination aspect.

Considerable research has been reported on reduced anxiety in group exams compared with individual exams (Caldecott & Emmioglu, 2017; Levy et al., 2018; Siegel et al., 2015; Vázquez-García, 2018). Although not directly addressed in this study, the students responded in the survey that they "felt comfortable receiving individual questions in the presence of the other group members" (Figure 2, Q3) and that they "experienced support in receiving and answering questions as a group" (Figure 2, Q2), which may suggest that the students experienced a low level of anxiety. This may explain, in part, the positive manner in which the students have embraced this oral group examination method with which they have minor or no previous experience.

Interestingly, the Biotechnological Engineering students and the second-year students agreed more on experiencing a fair exam in comparison with the Chemical Engineering students and the third-year students, respectively. The amount of group work presented to the Biotechnological and Chemical Engineering students is the same during their semesters. However, only the Biotechnological Engineering students have previously attended a group exam, whereas the Chemical Engineering students have not. The group exam tried out by the Biotechnological Engineering students was in their second-year, just prior to the present course. This most likely reflects the importance of becoming accustomed to a new examination method (Kolmos & Holgaard, 2009).

To match expectations between students and examiners, the authors strongly recommend carrying out a trial exam with the students prior to implementing this highly structured oral group examination method. During the trial exam, the skill of rephrasing questions and providing hints instead of answers should be demonstrated and trained with the students. Rephrasing questions and providing hints instead of answers informs the examiners that other students know the answer, but leaves room for the student being questioned, to come up with the right answer. In addition, the authors recommend that the students are trained to raise their hand to indicate a known answer. A simple indication provides time for all students to consider the answer and emphasizes that the exam is not a competition among group members. Moreover, to support a shared focus during the individual examination part, only one individual question should be revealed at a time, which was also explicitly noted by nine students. In practice, this may be orchestrated by letting all students randomly select a number for the examination instead of a question from a pool of questions covering the statistical methods taught during the course. When it is time for the individual student to be examined, the examiner reveals the actual question corresponding to the selected number. Finally, the authors recommend having the same number of students in all groups in order to facilitate an optimal flow during the examination of several groups. Group sizes of three to five are manageable (Billings, 2017). However, the authors prefer four students per group. Four students support group discussions better than three and are easier to manage than five when assessing the performance of each individual student.

In the present study, second year students were randomly assigned into groups, while third year students were self-assigned. It is worth noting that the survey did not reveal any significant difference between second and third year students concerning their level of feeling supported while being examined as a group (Figure 2, Q2). Interestingly, Billings has previously reported that self-assignment into groups positively influences group dynamics (Billings, 2017). No differences in grades were found in the present study between randomly assigned groups (second-year students) and self-assigned groups (third-year students), which was also reported by Nafziger *et. al*., who found no apparent differences in performance between randomly assigned and self-assigned groups (Nafziger, Meseke, & Meseke, 2011).

CONCLUSION

In conclusion, a highly structured oral group examination method was developed and implemented. Thus, an additional method has been added to the toolbox of oral examination methods. The examination method embraces both collaborative and individual assessment that aligns well with courses comprising both collaborative and individual learning and activities. Additionally, the group exam settings do not seem to hinder the students' perception of being able to perform individually. As the results in this study were based on a limited number of students that have tried the oral group examination method only once, it would be of great interest to perform a larger study, including more students, in order to compare and discuss in greater detail the efficacy of different oral examination methods that assess both collaborative and individual learning.

The greatest benefit from changing to the highly structured oral group examination method, however, is the extended examination time, which allows the examiners to assess the students in a more satisfactory manner. In addition, the highly structured oral group examination method provides the opportunity to address higher taxonomic levels of knowledge. As an added bonus, the need for the examiners to ask the same questions repeatedly is greatly reduced. While the highly structured oral group examination method does not require additional time-use per student for the examiners, it does increase the examination time for the students. However, compared with their previous project exams, where students are required to wait while each group member defends their group project individually, the increase in examination time is minimal. Another benefit of the examination method is the included evaluation intermission, which enables the examiners to discuss the performance of each student and adjust the remaining examination accordingly.

Based on the present study, we encourage other educators to explore this highly structured oral group examination method either *de novo* or to develop an existing individual oral examination method further.

ACKNOWLEDGMENTS

The authors thank the second- and third-year Chemical and Biotechnological Engineering students, Institute of Biotechnology and Chemistry, Aarhus University, Denmark, for filling out a questionnaire on their perception of the highly structured oral group examination method following their attendance in the course "Statistics and Design of Experiments (DoE)" in the spring semester of 2018. In addition, the authors thank Freia Martensen, Aarhus Institute of Biotechnology and Chemistry, Aarhus University, Denmark, for thorough proofreading of the manuscript.

REFERENCES

Bay, D., & Pacharn, P. (2017). Impact of group exams in a graduate intermediate accounting class. *Account. Educ., 26*(4), 316-334. doi:10.1080/09639284.2017.1292465

Billings, D. M. (2017). Collaborative Testing. *J. Contin. Educ. Nurs., 48*(7), 302-303. doi:10.3928/00220124-20170616-06

Bloom, B. S. (1956). *Taxonomy of educational objectives: the classification of educational goals*. New York: David McKay.

Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing : a revision of Bloom's Taxonomy of educational objectives* (Abridged ed.). New York: Longman.

Bretz, S. L. (2012). Navigating the Landscape of Assessment. *J. Chem. Educ., 89*(6), 689-691. doi:10.1021/ed3001045

Brown, S. A., & Glasner, A. (1999). *Assessment matters in higher education : choosing and using diverse approaches* (1 ed.). Buckingham.

Caldecott, M., & Emmioglu, E. (2017). Do Group Exams Support English as an Additional Language Student Learning? *Multicultural Learning and Teaching, 12*(1), 27-48. doi:10.1515/mlt-2014-0013

Hanna, K., Roberts, T., & Hurley, S. (2016). Collaborative Testing as NCLEX Enrichment. *Nurse Educator, 41*(4), 171-174. doi:10.1097/NNE.0000000000000241

Jang, H., Lasry, N., Miller, K., & Mazur, E. (2017). Collaborative exams: Cheating? Or learning? *American Journal of Physics, 85*(3), 223-227. doi:10.1119/1.4974744

Kolmos, A., & Holgaard, J. E. (2009). Gruppebaseret eller individuel projekteksamen – fordele og ulemper. *Dansk Universitetspædagogisk Tidsskrift, 4*(7), 33-43.

Krogh, L., & Aarup Jensen, A. (Eds.). (2013). *Visions, challenges, and strategies : PBL principles and methodologies in a Danish and global perspective* (1 ed.). Aalborg: Aalborg University Press.

LaBossiere, M. D., Dell, K. A., Sunjic, K., & Wantuch, G. A. (2016). Student perceptions of group examinations as a method of exam review in pharmacotherapeutics. *Currents in Pharmacy Teaching and Learning, 8*(3), 375-379. doi:10.1016/j.cptl.2016.02.002

Levine, R. E., Borges, N. J., Roman, B. J. B., Carchedi, L. R., Townsend, M. H., Cluver, J. S., . . . Thompson, B. M. (2018). High-Stakes Collaborative Testing: Why Not? *Teaching and Learning in Medicine, 30*(2), 133-140. doi:10.1080/10401334.2017.1365719

Levy, D., Svoronos, T., & Klinger, M. (2018). Two-stage examinations: Can examinations be more formative experiences? *Active Learning in Higher Education*, 146978741880166. doi:10.1177/1469787418801668

Lindsley, J. E., Morton, D. A., Pippitt, K., Lamb, S., & Colbert-Getz, J. M. (2016). The Two-Stage Examination: A Method to Assess Individual Competence and Collaborative Problem Solving in Medical Students. *Academic Medicine, 91*(10), 1384-1387. doi:10.1097/ACM.0000000000001185

LoGiudice, A. B., Pachai, A. A., & Kim, J. A. (2015). Testing together: When do students learn more through collaborative tests? *Scholarship of Teaching and Learning in Psychology, 1*(4), 377-389. doi:10.1037/stl0000041

Mahoney, J. W., & Harris-Reeves, B. (2017). The effects of collaborative testing on higher order thinking: Do the bright get brighter? *Active Learning in Higher Education*, 1-13. doi:10.1177/1469787417723243

Molsbee, C. P. (2013). Collaborative testing and mixed results. *Teaching and Learning in Nursing, 8*(1), 22-25. doi:10.1016/j.teln.2012.09.001

Momsen, J., Offerdahl, E., Kryjevskaia, M., Montplaisir, L., Anderson, E., & Grosz, N. (2013). Using Assessments to Investigate and Compare the Nature of Learning in Undergraduate Science Courses. *CBE-Life Sciences Education, 12*(2), 239-249. doi:10.1187/cbe.12-08-0130

Nafziger, R., Meseke, J. K., & Meseke, C. A. (2011). Collaborative testing: the effect of group formation process on overall student performance. *The Journal of chiropractic education, 25*(1), 11-15. doi:10.7899/1042-5055-25.1.11

Pienta, N. J. (2011). Striking a Balance with Assessment. *J. Chem. Educ., 88*(9), 1199-1200. doi:10.1021/ed200442j

Rhodes, M. M. (2010). A Laboratory Practical Exam for High School Chemistry. *J. Chem. Educ., 87*(6), 613-615. doi:10.1021/ed100200k

Rieger, G. W., & Heiner, C. E. (2014). Examinations That Support Collaborative Learning: The Students' Perspective. *Journal of College Science Teaching, 43*(4), 41-47. doi:10.2505/4/jcst14_043_04_41

Rivaz, M., Momennasab, M., & Shokrollahi, P. (2015). Effect of collaborative testing on learning and retention of course content in nursing students. *Journal of advances in medical education & professionalism, 3*(4), 178-182.

Siegel, M. A., Roberts, T. M., Freyermuth, S. K., Witzig, S. B., & Izci, K. (2015). Aligning Assessment to Instruction: Collaborative Group Testing in Large-Enrollment Science Classes. *J. Coll. Sci. Teach., 44*(6), 74-82.

Vázquez-García, M. (2018). Collaborative-group testing improves learning and knowledge retention of human physiology topics in second-year medical students. *Adv. Physiol. Educ., 42*(2), 232-239. doi:10.1152/advan.00113.2017

Wieman, C. E., Rieger, G. W., & Heiner, C. E. (2014). Physics Exams that Promote Collaborative Learning. *The Physics Teacher, 52*(1), 51-53. doi:10.1119/1.4849159

BIOGRAPHICAL INFORMATION

Dr. ***Karin Stenderup*** is an Associate Professor at Institute of Biotechnology and Chemistry, Faculty of Technical Sciences, Aarhus University, Denmark. She holds an MSc degree (1999) in Chemistry and Biotechnology and a Ph.D. degree (2002) in Molecular and Structural Biology from Aarhus University. Karin has been working professionally with teaching in higher education at Institute of Biotechnology and Chemistry, Aarhus University since 2015.

Dr. ***Sanne Sandberg Overby*** is an Associate Professor at Institute of Biotechnology and Chemistry, Faculty of Technical Sciences, Aarhus University, Denmark. She holds an MSc degree (2002) in Biotechnology from Aalborg University. Sanne has interests in didactics and dissemination and has experience with education and teaching development.

Corresponding author

Karin Stenderup
Institute of Biotechnology and Chemistry
Faculty of Technical Sciences
Aarhus University
Hangoevej 2
DK-8200 Aarhus N
Denmark
Karin.Stenderup@bce.au.dk

SUPPLEMENTARY MATERIAL

A “ready-to-use" Student Scoring Sheet to ease the task of keeping track of each individual student’s contribution during the examination. (An example is provided in *red*)

Student Scoring Sheet

(A quick and easy way to document the individual student’s contribution during examination)

Group: _____*3*_____

1. Group examination (based on the handed-in report)[1]:

Learning objectives	Student 1[2]	*John*	Student 3	Student 4	Report notes
Path of steepest ascent	*2*	*-1*			
Alias structure			*0*	*1*	*0-1?*
.....					
.....					
Add as many as needed					

2. Individual focus:

	Student 1	Student 2	Student 3	Student 4
	Chosen question 1[2]	*ANOVA*	Chosen question 3	Chosen question 4
Time[3]	Start ____ Finish ____	Start *0:00* Finish *5:59*	Start ____ Finish ____	Start ____ Finish ____
Excellent answers to:		*SST, SSA, SSW*		
Right answers to:		*Assumptions*		
No answers to:		*RE*		
Wrong answers to:		*Two-way*		

3. Group reflection:

	Student 1	Student 2	Student 3	Student 4
Additional contributions or comments?		*Q1: right* *Q2: wrong*		

Notes:

1. To keep track of each student’s contribution during the group examination, a score is given from -1 to 2, where -1 is given for “wrong answer”; 0 for “no answer, 1 for “right answer”, and 2 for “excellent answer”.
2. At the examination, actual student names and chosen questions are recorded on the sheet.
3. Start and finish times are recorded to ensure that each student is examined in the individually allocated time.