
AC 2012-3459: A STUDY ON THE EFFECTIVENESS OF TEAM-BASED ORAL EXAMINATIONS IN AN UNDERGRADUATE ENGINEERING COURSE

Mrs. Lisa K. Davids, Embry-Riddle Aeronautical University, Daytona Beach

Lisa Davids is an Associate Professor in the Freshmen Engineering Department, having taught at Embry-Riddle Aeronautical University (ERAU) for 12 years. She is currently the Program Coordinator for the Freshman Engineering Department and serves as the course monitor and coordinator for the Introduction to Engineering course. During her tenure at ERAU, she has taught Fluid Mechanics, Dynamics, Experimental Aerodynamics, Aerodynamics I, and Introduction to Engineering. She is the Faculty Advisor for the ERAU chapter of Society of Women Engineers as well as a Co-advisor for the all Women's Baja SAE Team at ERAU. Her research interests involve the retention of women in engineering degree programs and effective pedagogy in undergraduate engineering curriculum.

A Study on the Effectiveness of Team-Based Oral Examinations in an Undergraduate Engineering Course

Abstract

The conventional and pervasive written exam format used in undergraduate engineering courses, while practical, may be neither the most effective assessment tool in determining a student's ability to perform in the professional arena, nor an effective tool to motivate students to understand the fundamental concepts on which the exam is based. The effectiveness and practicality of employing team-based oral examinations in lieu of individual written midterm exams in an undergraduate course will be presented in this paper. Several of the disadvantages of written exams are addressed with the protocol developed for the oral exams. These protocols are defined and discussed. Aggregate student performance on the oral exams and the written final exam from the experimental group are presented and examined. Effectiveness of the team-based oral exams is quantitatively assessed through a comparison of student performance between the control (individual written exams) and the experimental group (team-based oral exams). The final exam average for the two groups is provided as a point of comparison to determine assessment impact. Additionally, the effectiveness of the team-based oral exams from the students' perspective is quantitatively and qualitatively assessed using results from a 5-point Likert survey and candid comments. The subjective experiences from the student and professor perspectives are presented. Survey results indicate an increase in the effort, retention and recall of knowledge of the students, a finding also supported by the final exam average comparison. The team-based oral exams appear to serve as a more effective tool than the traditional testing method to evaluate students' learning and facilitate their deeper understanding of the course content.

Reasoning

The typical technical written exam (in STEM courses) is the most common way to assess a student's learning when the class sizes are large, or the professor has multiple sections. However, it may not be the most accurate assessment technique in general and certainly is not reflective of how the students will be expected to perform once they become professionals. Never will a student be expected to create a report of analysis for a well-defined problem without external resources or help, without peer review, in an hour, only to wait for someone to review their report without any opportunity to further explain or defend their analyses. Not only does this provide the professor an incomplete picture of the students' skills and understanding, but if replied upon as the sole assessment throughout their undergraduate tenure, it would create students who are ill-prepared to orally communicate their ideas and analyses, as they will have to in the professional world. The ability to communicate and defend one's position is a requirement of the successful engineer^{1,2} and unfortunately has become a skill not well-developed in today's engineering undergraduates³. Additionally, collaboration and working within a team environment is another essential skill for the practicing engineer, and hence engineering student^{2,4}.

A one hour written exam limits the assessment to the students' ability to perform only in that one hour of time. If the student had a difficult day, poor night's sleep or any other wealth of personal issues, the impact on their performance during that hour could likely be significant⁵. A one hour written exam also limits the complexity of the problems. With only an hour, it is difficult to give students problems beyond a certain level of difficulty or complexity. A problem worth solving should require some time to reflect and the consideration of a few different approaches. One hour is simply not enough time to allow for this without requiring the student to perform quickly; speed should not be a factor in the testing process^{6,7} when effectively evaluating students' knowledge.

To attempt to reconcile these shortcomings, oral examinations were used in two sections of an undergraduate engineering course (Introduction to Fluid Mechanics). Oral examinations, used for most graduate and post-graduate programs to assess whether or not a student is qualified to continue in the program, are well known to be effective in determining a student's level of understanding⁸. However, the time-intensive nature of individual oral examinations is perceived to be prohibitive of their use in today's undergraduate class sizes and faculty work loads. To address this issue, an innovative technique of team-based oral examinations was attempted. The benefits of an oral defense of a student's solution, peer review and team collaboration, and use of resources will all be recognized, while maintaining a reasonable burden of time when assessing a team of students rather than individuals. This technique is similar to team-based design report presentations, with the differences in the level of open-endedness and size of the audience. Unlike project presentations, the oral examinations were presented to a closed audience comprised of the professor and the students' team. Although this is less reflective of the professional world, as this is an introductory course, the objective for these exams was to assess their knowledge, not yet their ability to handle the stress associated with performing in front of a full audience of their peers. Assessment of student knowledge should be performed in a setting that minimizes test anxiety⁹. Although oral examinations can be stressful, the team-based support and the fact that there is no other audience should minimize the stress associated with oral exams. Additionally, the stress associated with written exams has been shown to be detrimental for some students¹⁰.

The purpose of this pilot study was to determine the feasibility of conducting team-based oral exams in an undergraduate course, both in terms of the time burden on the instructor as well as performance of the student. The main concern of the investigator was the question of individual accountability in a team setting. Given the format outlined in this paper, can an individual's knowledge be assessed when the solutions were developed as a team? Although this question is not definitively answered, the results of this pilot study seem to support that with refinement, this technique of examination can be effective at individual assessment. It was also found that once the questions are formulated for the exams, the time burden on the instructor was not significantly more when compared to that for the conventional exam. The pilot study was successful in verifying that this method of examination seems to be feasible and warrants further and more focused investigation. The next steps in future studies will involve the determination of the accuracy and validity of this assessment technique.

Procedure

Preparation

During the first day of class of the experimental semester (spring 2011), the students were informed of the experimental testing policy so they could make an informed decision on whether or not to stay in the section (several other sections of the same course were available at the same times with open seats). Students were made aware that remaining in the section was an agreement to participate in the study. Though official IRB approval was not obtained (this study was initialized as an informal pilot study for which no funding was received), informed consent was obtained from each student. IRB approval will be obtained prior to all future studies.

A table of timeslots (availability table) for the dates of the two scheduled examinations was provided for the students to complete. Students were asked to mark the time slots during which they were not available for an oral examination. An hour was allowed for each time slot with a half hour between each time slot in case of overage and also time for a break before the next team. The number of time slots exceeded the number of total teams anticipated for all sections. The time slots were limited to the minimum number of days possible to maintain as much equality in preparation time for the teams as possible.

The student feedback on the availability table was used to ensure all timeslots were useful or to change/add timeslots as needed, and also to facilitate effective teaming for the students. The data was compiled and shared with the students so that they could choose teammates that shared oral examination timeslot availability.

Teams were formed as soon as possible so that the team members could become acquainted and become familiar with their individual team dynamic while working on homework assignments. This was meant to better prepare them to collaborate and perform as a team for the examinations. Although the intent was to form the teams by the third week of classes, the process involved in reducing the data from the availability tables required far more time than expected. Teams were not formed until the week that the first midterm exam was administered (the fifth full week of the semester).

Teams

The teams consisted of four students and were self-formed by the students. However, the team formation was facilitated by the professor and the information from the availability tables. Students were grouped into subsets of 12 based on three shared timeslots for which all 12 students were available. From that group of 12, students chose their teams of four. Once the students knew the make-up of their team, they were able to reserve one of the three timeslots assigned to their subset. Reservations were taken on a first come/first served basis. To allow for optimal teaming and use of time, teaming across the two sections was permitted. To maintain alignment with this allowance, no team-based activities were employed in class.

Execution of the team based oral exams

The three sections taught by the investigator during the spring 2010 served as the control group; the two sections taught by the same investigator during the spring 2011 served as the experimental group. Two midterm exams were administered during the control and experimental semesters. For both groups (control and experimental), the first midterm exam occurred during the sixth week of classes while the second midterm exam occurred during the 12th week of classes. (Both the control and experimental semesters were 14 weeks in duration). The two midterm exams were administered in the conventional written format for the control group while for the experimental group, the two midterms were given in the team-based oral format.

For the control group, the written exams each consisted of four fairly simple problems which covered the material only on which students had turned in homework. The written exams were administered during the 75 minute class-session and allowed 1 equation sheet. For the team-based oral format, each team of four students was provided a team-unique set of problems. Each problem set contained four problems (equal to the maximum number of students in a team), and again, only covered the material on which students had turned in homework. The teams were given their problem sets in class the week prior to the oral examination timeslots. Some teams had exactly 1 week, others had a little more than a week to prepare for the oral examination, depending on the assigned oral exam timeslot.

The entire team was expected to be prepared to discuss each of the four problems on the test; however each student was responsible for presenting and defending at least one problem during the oral exam. Students were not informed prior to the oral exam time which problem they would be asked to explain; in fact, the problems were randomly assigned at the beginning of the one hour oral exam. This measure was used to help ensure the students would study and be prepared for all of the problems in their set.

The oral exam rubric (see Appendix A) was developed to assess the students' knowledge during the oral exam. Each student received their individual oral exam grade based on the validity of their solution and their ability to present their solution and defend it under scrutiny from the professor. The rubric was supplied and explained to the students prior to the first exam. The rubric includes seven categories on which the students were assessed. The rating scale is also defined in terms of performance attributes during the oral exam. Ratings were determined during (or in some cases, immediately after) the individual student's presentation to ensure the most accurate assessment.

During the experimental semester, only the two midterm exams followed the team-based oral exam format. The final exam consisted of a typical written format for new material not previously tested and a weighted multiple choice format for the cumulative portion. The final exams used in both control and experimental groups were nearly identical, changed only where needed to minimize opportunity for answer sharing. It was this final written exam that was used as the litmus test to gauge the feasibility and effectiveness of the team-based oral examinations.

Course Grade Determination

Due to the uncertainty in validity of the oral exam method of assessment, each oral exam midterm was weighted as 20% of the students' overall grade. This was lower than the usual 30% for each midterm that was used in previous semesters. The final exam and quizzes were both increased (final – from 20% to 35%, quizzes – from 10% to 15%) to compensate for the reduced midterm weights.

Written exams are generally more objective than oral exams; however, oral exams should provide a more realistic mode for students to demonstrate their knowledge, thus allowing a more accurate assessment of that knowledge. As a baseline, the final exam was retained as an individual written exam and shared many questions with final exams used in past years. The similarity in format and level of difficulty of the final exam provides a point of comparison in the students' assessment for validation of this oral exam process. Based on previous semesters, the past final exam average has been in the low 'C' range, validating the level of difficulty of the final exam; a similar trend should be expected if the oral examination method is as effective as the traditional evaluation method.

Results

Oral Examination – Student Performance Data

The data shown in the boxplot (Figure 1) demonstrate a large difference in the median scores between the two oral midterms when compared to the written final exam for the experimental group (all shown in solid blue). The median grade for the first midterm was a 93 and the median for the second midterm was a 91.5. Compare these to the written final exam median of 76. The sample size remained constant throughout the semester at 76 students (no students dropped the course). Clearly, on average, the students performed at a much higher level on the oral exams when compared to the written final exam.

For the control group, the first midterm median grade was a 78, the second midterm median grade was an 84 and the final exam median was a 70. The sample size for these averages varied from 99 to 93 as students dropped the course throughout the semester. Grades of zero (from the dropped students) were not included in the medians.

The most dramatic difference between the two semesters is the distribution of grades. The grades from the control group semester cover a much broader range. The standard deviation from the average grade for both midterms in the experimental group was 10 percentage points (E1-exp = 9.6, E2-exp = 10.4). For the control group, the standard deviations for the two midterms were 15.4 and 12.1 (E1-cont and E2-cont, respectively). However, when comparing the final exam grades, the control group showed a slightly smaller standard deviation when compared to the experimental group; the standard deviation for the experimental group final exam grades was 14.7 while the control group standard deviation on the final exam was 13.8.

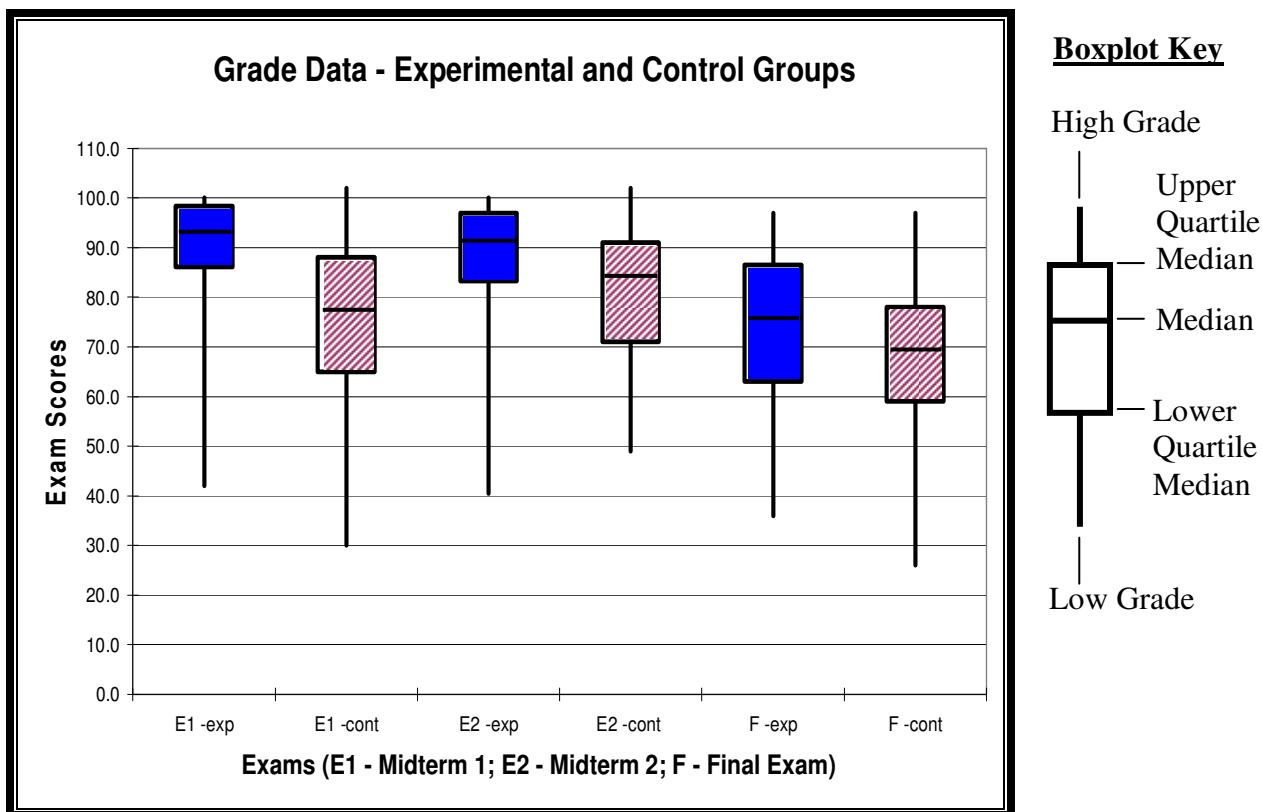


Figure 1: The above figure shows the boxplots for both the experimental group (spring 2011) and control group (spring 2010). The two midterms (E1–exp and E2–exp) were administered using the team-based oral format while the final exam (F–exp) was administered in a written format. The experimental data is shown in solid blue. The control group exams were all in written format and are shown in the cross-hatched red.

When the actual grade distribution for each semester is considered, the benefit of the oral exam format is apparent. The percentage of grades earned on each of the exams (both midterms and the final) is broken out and provided in Figures 2 and 3. The most impactful data is the final exam data (represented by the third bar in each grade grouping). The final exams were nearly identical between the two semesters (only minor differences exist between the exams to account for multiple sections and the attempt to minimize any opportunity to share exam answers between final exam times). Though both semesters share a similar “shape” in grade distribution, the experimental group shows a marked increase in grades. A larger percentage of students from the experimental group earned A and B grades on the final exam, and a smaller group earned D and F grades.

The percentage increase in A grades from the control group to experimental group was 75% (control – 8%, experimental – 14%) and for B grades the percent increase was 31% (control – 16%, experimental – 21%). Commensurate with the increase for the A and B grades, there was a corresponding decrease in the D and F grades. The percent decrease in D grades from the control group to experimental group was 14% (control – 14%, experimental – 12%) and for F grades, the percent decrease was 30% (control – 30%, experimental – 21%). Interestingly, the percentage of students earning a C on the final exam was equal between the two groups (32%).

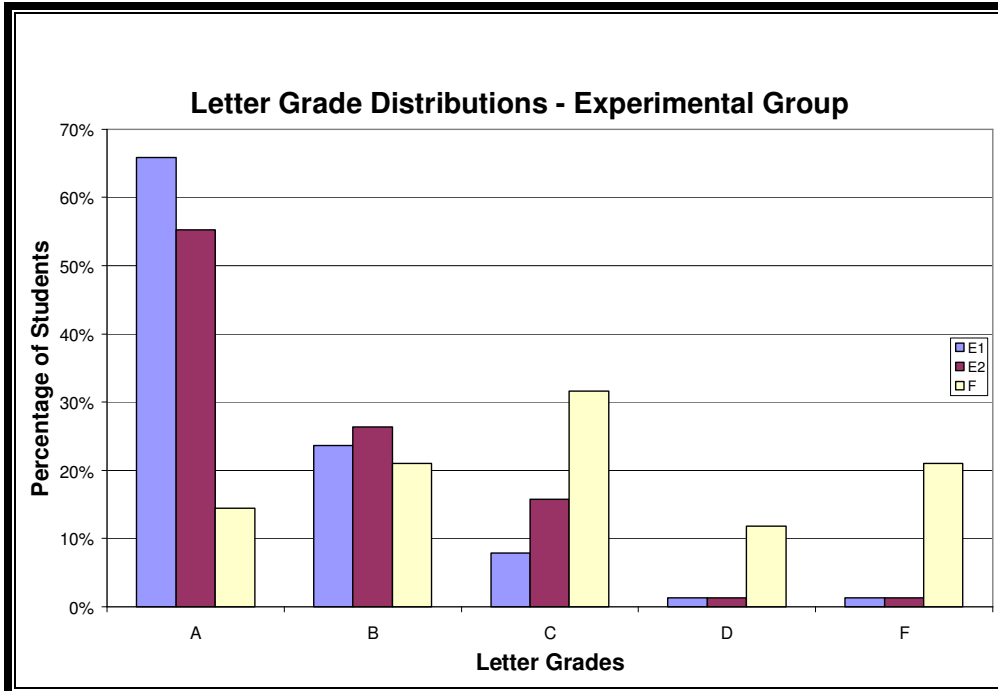


Figure 2: This figure shows the distributions of letter grades for each exam for the experimental group (spring 2011). The first bar in each letter grade groups represents the percentage of students earning that grade on the first midterm; the second bar represents the second midterm, and the third bar represents the final exam. The impact of the oral midterms appears to be in the overall success of student performance on the final exam.

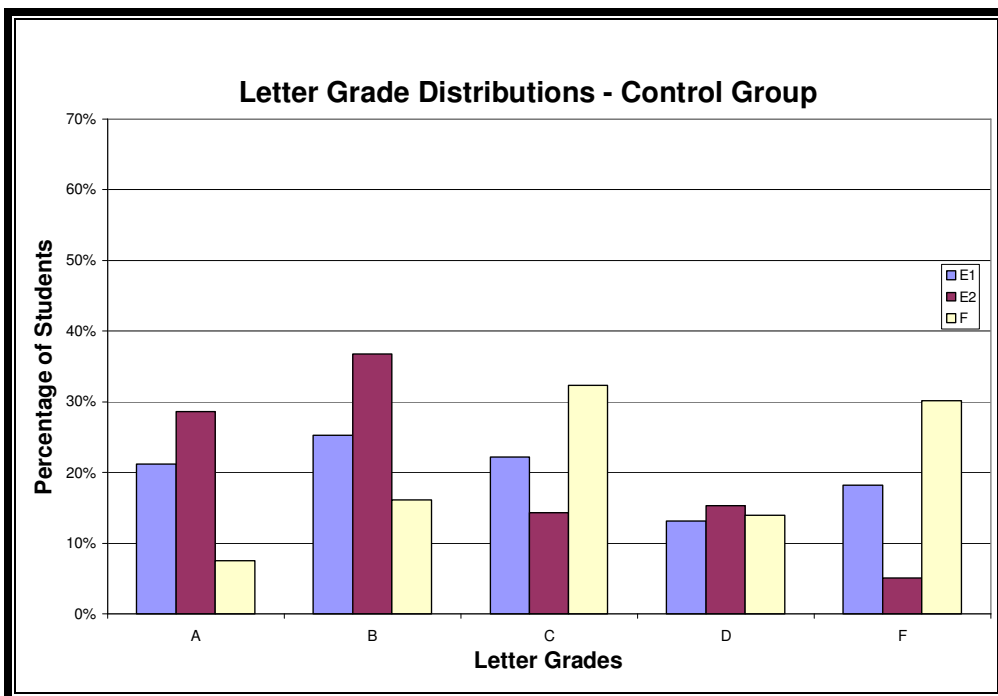


Figure 3: The distribution of grades is once again shown in the above figure; however, these are the data from the control group (spring 2010).

The higher final exam average in the experimental group with the slightly larger standard deviation results in the higher number of A and B grades when compared to the control group. It is felt that this demonstrates support in favor of the oral exam method of student assessment. The overall performance from both the control and experimental semesters are provided in Figure 4 for a summarized comparison of the students' average performance. As can be seen in this figure, the average grades were higher for the experimental group, including the common final exam.

The higher midterm exam scores for the experimental group and the smaller standard deviation values can likely be explained by a few different factors: 1) the week of preparation and team-based environment resulted in a substantial decrease in tedious mistakes (such as in arithmetic and units), 2) luck was not a major factor in student performance as they knew that they would be required to defend one of the four questions provided for which they had a week to prepare; this can increase student expectations of their performance (confidence) and hence, actual performance¹¹, 3) Students preparation was more in depth as they knew exactly where to focus their efforts, 4) the newness of the oral exam (real-time) grading experience for the instructor may have caused a more lenient grading effort.

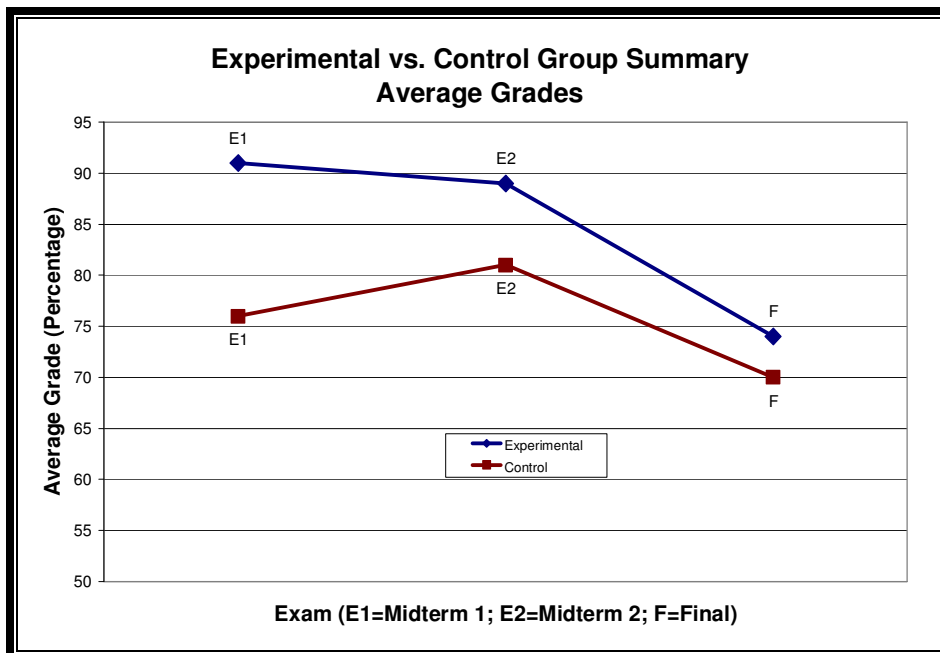


Figure 4: The above figure compares the average grades for the two midterm and final exams between the control and experimental groups. The control group utilized the conventional written exam format for all exams; the experimental group utilized team-based oral exams for the two midterms.

Limitations

The primary limitation of this study was the year lag between the control group and experimental group. Ideally, one section from the same semester should have been used as the control and the

other for the experimental to ensure students received as identical as possible instruction. However, as students talk to one another between sections, the newness of the examination method in the would-be experimental group may have caused self-selection of students into their “preferred” section thereby affecting results. It was determined for fairness to the students and to prevent a self-selected control or experimental group, both sections of the spring 2011 term were part of the experimental group and the control group was taken as the previous spring 2010 semester. Related to this limitation is the possible difference in skill level of the students. However, as both the control group and experimental group occurred in their respective spring semesters (control – spring 2010 and experimental – spring 2011), the students should have been at the same point in their curriculum in either group and hence their general skill level should have had no reason for a significant difference. The percentage of students repeating the course was not compared between the groups as this data was no longer available for the control group at the time of the writing of this paper.

A third limitation involves the reliability of the experimental data. The reliability of the control data could be easily verified by considering several past semesters with the use of written exams; however, until the oral examination technique is used for several semesters, the reliability of the experimental data remains unverified. Finally, the validity of the oral examination technique is being determined based on how the experimental group’s performance compared to the control group’s performance on a written final exam and not the oral exams themselves. As the uncertain validity of the written exam has been the impetus of this study, using it as the baseline and comparison point is limiting. However, the difference in average grades for the two midterm exams (between the control and experimental groups) is difficult to resolve when the accuracy and/or validity of the control method is called into question. A separate method to gauge the accuracy and validity of either examination method is needed. Focusing on validating the team-based oral exam method without the need to compare to the written exam method will be the direction of future studies.

A fourth limitation on the study is related to the first and third limitation; namely, that due to the nature of the experimental exam format, the level of difficulty in the questions for the team-based oral exams was much higher than those of the control group. The students were allowed to work in teams for a week on the solutions to the oral exam problems. To ensure that the students were appropriately challenged, the questions were chosen to require the transfer of knowledge to diverse contexts and the synthesis of component skills. Given the difference in level of difficulty, the comparison between the control and experimental midterm exams is limited. Additionally, the difference in the median and average grades between the experimental midterms and final exam is difficult to explain. Are the midterm grades higher due to the improved learning environment, or for some other reason? Which exam method provides the more accurate assessment, the oral exam or the written exam? This limitation will be avoided in future studies with the focus shifted to validation of the team-based oral exam technique without comparison to the conventional technique. For the pilot study, this comparison provided sufficient data for gauging feasibility and providing future direction for more focused studies.

Survey

For the initial implementation of this technique, student feedback was used to evaluate its effectiveness. A feedback survey, provided in Appendix B, was given to the students following the first oral examination. Figure 5 illustrates the response percentages for each survey question. Questions 1, 2, and 5 achieved 100% positive feedback (or agreement). Questions 1 and 2 focused on how the experimental testing procedure (the week long preparation for the oral exam format) affected their learning and understanding of the material. Question 1 asked: “Compared to conventional testing, the testing procedure used for this first exam facilitates broader learning for me.” Question 2 asked: “Compared to conventional testing, this method of testing facilitates more in-depth learning on the particular concepts that related to my questions.” Question 5 relates strictly to the oral examination portion of the exam and asked: “By having to explain my solution in person, I feel this method better prepares me for the professional working environment.”

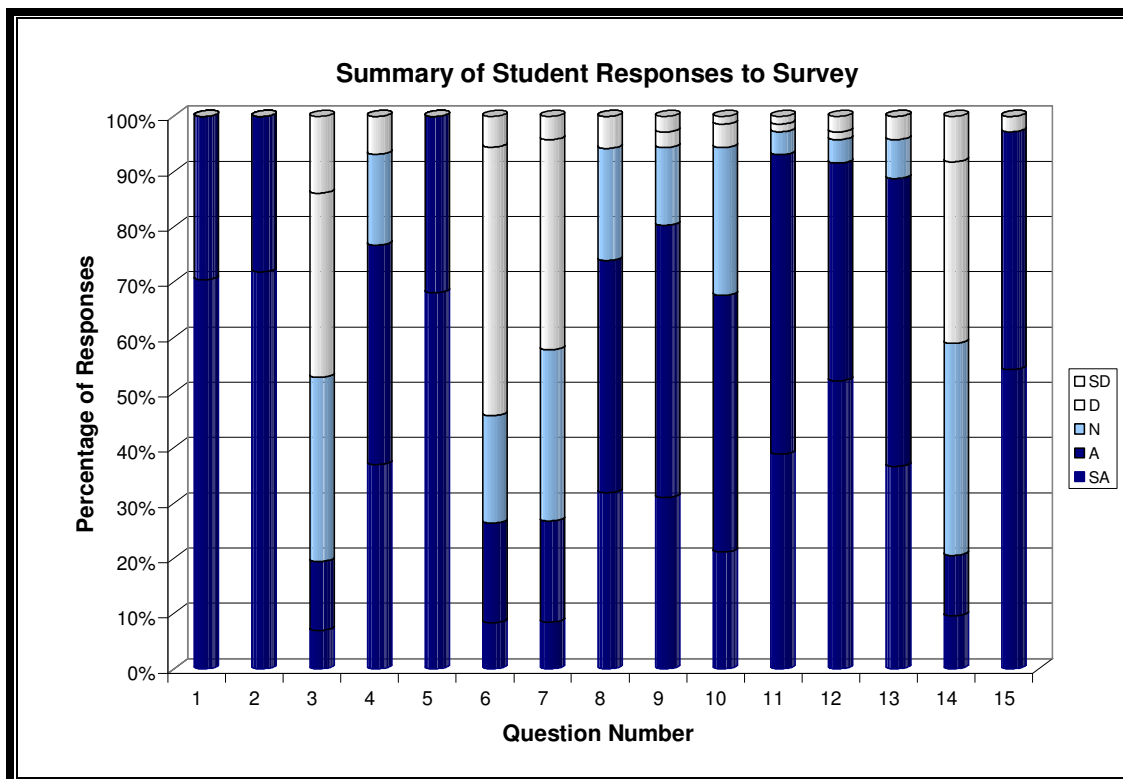


Figure 5: The above figure summarizes the response percentages for each survey question. For simplicity, the Strongly Agree and Agree responses are shown with the same dark hue, the Neutral responses has a light shade, and both Disagree and Strongly Disagree are shown in white. The data is displayed in the column graphs in the same order as the key on the right.

High levels of agreement (over 80%) were achieved on questions 11, 12, 13, and 15, which all focused on the logistics of the testing procedure (team-size, preparation time, level of difficulty and oral exam time). About three-fourths of the students agreed that this method testing allows for a more accurate reflection of the students’ knowledge (question 4). Questions 3, 6, 7 were worded such that a response of disagree or strongly disagree showed support for the oral exam or

the teaming environment used; question 14 inquired on whether the oral exam portion would be better completed individually rather than in the team setting.

Based on these survey results and end of semester evaluation comments provided by the students, this method of testing is viewed more favorably than the typical written exam. Several students noted that they put more time and effort into their preparation for these two oral exams than they did any other exam. They appreciated and took advantage of the ability to check several resources for approaches, recheck their solutions, retry their solutions, brainstorm and verify ideas with their team, and study the specific concepts related to the problems in preparation for the oral defense. Knowing that they had to be able to not only explain what they did and why but that they may have to answer spontaneous questions that would probe true understanding motivated them to learn the material at deeper level than in their other courses. It is this longer and deeper engagement with the material that leads the investigator to believe that a higher level of mastery is facilitated through this method of examination. One quote reads: “Not only was this type of exam less stressful but most importantly I gained a better understanding of the material. At the end of the semester I can still recall all of the concepts from the first two exams in this class where I have trouble doing so for the other classes that use the standard written test. I also feel that more of my knowledge in the course material was tested than with a conventional written test. In addition to a better understanding of the material, I feel that this test closely mimicked a real world scenario, in which you would be a part of a team working on an engineering project.” Another quote reads: “I really liked the take home test method. The increased difficulty of the test forced me to do plenty of research and trial and error, which in turn helped me learn while doing the test. I also feel that it better tested my knowledge of the material, as I was able to show me understanding of all the principles and concepts without worry of blanking on one little formula.” Two quotes commenting on the amount of time dedicated to preparation for the exam read: “I spent so much more time preparing for the oral exams, seven or eight hours, but it was so much more worth it” and “A lot of work had to be put into preparing for the exam sessions, rather than just studying and hoping to not have forgotten something. When part of the work was incorrect, or was not quite right, talking about it and being prodded to realize the correction was much more useful than just seeing it marked wrong on paper”. Additional quotes are provided in Appendix C.

These quotes touch upon nearly every aspect essential for learning to occur. For example, students seem to be more motivated by this method of examination. They see the value and have more favorable expectancies in their ability as the uncontrollable factors have been removed (such as luck). This method also seems to facilitate a higher level of mastery. The challenge of the questions require synthesis of component skills as well as transfer of knowledge, as the questions call upon knowledge and skills acquired from pre-requisite courses as well as from the current course. These mastery skills are facilitated by the team-based nature of the exam as well as the unlimited resources and time allowance. These conditions result in the students’ longer and more in-depth engagement with the material. Finally, the real-time feedback provided during the oral examination session provides both the instructor and student the chance to learn where the students’ deficits in understanding are, creating the most productive atmosphere for learning.

Professor Perspective

Concerns

The main concern that surfaced both from students and the instructor was the potential inconsistency in the level of difficulty of the questions – both across “teams” and within the same team. Some teams’ exams were “overall” more difficult than others. Within any given team, there was almost always one very difficult problem and one fairly simple problem. This made a fair and representative assessment difficult. The obvious solution is to use problems with more consistent levels of difficulty; however, execution of that solution is another issue. To create enough variation in the test problems, ensuring one student doesn’t solve the whole exam and all 75 other students copy, each exam needs to be unique. For two sections of 38 students each, this translates to needing 76 different questions, all of the same level of difficulty. That is quite a daunting task and should be accomplished prior to the start of the term.

As with a written exam, there is still a concern that the instructor is only evaluating a fraction of the picture of the individual’s knowledge as each of the students only talk about 25-30% of the exam. Since the oral exam portion was carried out in the team environment it was difficult to ask questions to two or more students on the same question, unless the first student gives an incorrect solution or reasoning. For this reason, moving to 20 minute individual oral exams (in lieu of the hour long team oral) might be more effective. To prevent answer sharing or other academic integrity concerns, it would be necessary to schedule the individual orals for each team member to be in succession and require that the team wait in a sequestered room.

Related to the last issue, there was also a concern about “question” sharing; teams of students discussed what spontaneous questions were asked during the oral interviews. This was a serious concern, especially when some of the questions were the same or very similar between teams. Some questions were very general, but others were unique or specific to the question and were used to probe a little deeper into the students’ understanding. This was effective only if the question was unanticipated.

There was also a concern about the consistency in the instructor’s assessment. As the instructor became more comfortable with the process, and with the real-time evaluation of the students, she felt she may have been more critical of the later teams (days 3 and 4). Additionally there was concern with the feedback to the students. The instructor’s conventional written exams provide a wealth of comments, corrections, and explanations so that the students can see what they did incorrectly and learn from their mistakes. With the timeframe of this exam, the instructor did not make any comments or corrections on the students’ work; it was strictly an assessment event. The reasons for this concern are trifold: 1) the instructor is accustomed to using exams not only as an assessment tool, but also as a learning tool. 2) For ABET purposes (and own personal satisfaction in the validity of the assessment), it is useful to have the assessment of the students documented. Having a clear rubric and written evidence of how that rubric was applied serves that purpose, but recording the interviews (which was not performed) would also serve that purpose. 3) Should a student disagree with the scoring, it is very difficult to go back and justify the reasoning for the evaluation without much evidence (if the interview is not recorded). The students were required to turn in their solutions, but they were not graded only on their written

solution, but also their explanation and perceived understanding of that explanation. With no recording, the evaluation would be difficult to recreate and validate. Clearly, there are solutions to this concern such as recording the interviews, taking generous notes and writing comments in real time.

Lastly, the attempt to interpret whether the student was simply nervous or truly lacked understanding of the problem was an issue. This issue came up with at least two students. They were so nervous before even beginning their interview that it was unclear how many of their mistakes were caused by lack of understanding versus overwhelming nervousness. This is also a factor with written exams, so it is not a unique problem to oral exams.

Lessons Learned

It is important to be very clear on how the students will be assessed ahead of time. The students were requested to write up their solutions so they would have something to which they could refer during the oral interview; however, it was not clear to the students if they should each write up their own solution, how neat or organized that solution should be, or whether or not it was even fully required. As a result, some solutions were written very poorly, disorganized, incomplete, or not at all. This made the assessment more lengthy and difficult. It should also be made clear that they will be graded both on what they have written at the time of the exam and what they say. If there is a mistake on the paper, it will count towards their grade and if there is a mistake said during the interview, it also counts.

It is important to have worked out every single solution ahead of time yourself. There was one example of a student who approached the problem incorrectly but explained it so well the mistake was not recognized until the instructor went through the analysis herself. This would not have happened had the analysis been completed ahead of time.

It would be helpful to record which team has which problems for preparation just prior to the interview. This includes ensuring that the solution is complete and reviewed so it is fresh in the instructor's memory.

Having the students complete a self and peer evaluation after the oral examination could be useful. The self evaluation could reflect the students' perception of their own knowledge and the peer evaluation could reflect their team's teamwork level. The self assessment need not be used to determine their grade but only used to recognize any problems either in the assessment or teaming procedures (i.e., if there is a large discrepancy between the professor's assessment and the students' self assessment of their knowledge, the professor can request a one on one conference to mitigate the discrepancy). Peer evaluations could be used to modify the teams for the future exams as needed.

While being an extremely effective educational tool, employing oral exams in a large undergraduate course proved to be a daunting and time-intensive venture, at least for the test creation phase. However, the results were very satisfying and the student feedback was very positive.

Future Work

In addition to continuing to utilize this oral exam format and refining the logistics of the examination, it would be useful to follow a cohort of students from the studied course into related higher level courses and compare their performances in those courses to students who were evaluated using conventional written exams in the original course with a different professor. This might provide insight into the validity of the oral exam grade results and claims made by the students that they learned the material more fully. Future studies will focus on validating the team-based oral exam method without relying on comparison to the conventional written exam method. A proper concept inventory for fluid mechanics will be researched and adopted to aid in future studies. Additionally, the ability of students to 1) transfer knowledge and 2) synthesize component skills will also be studied. Based on the student survey feedback, it is presumed that the additional level of challenge from the team-based oral exams required students to demonstrate both of these attributes of mastery. Though it is speculative at this point to proclaim the team-based oral exam method facilitated a higher level of mastery, the results of this pilot study indicate that a more in-depth look is warranted.

References

-
- ¹ James Harris, Eugene DeLoatch, et. Al., Journal of Engineering Education Round Table: Reflections on the Grinter Report, (JEE January 1994), pg 71, 72, 74 and 83
 - ² Robin Adams et. Al., Multiple Perspectives on Engaging Future Engineering, (JEE January 2011, Vol. 100, No. 1), pg 62
 - ³ L.L. Bucciarelli, H.H. Einstein, P.T. Terenzini, A.D. Walser, ECSEL/MIT Engineering Education Workshop '99: A Report with Recommendations (JEE April 200), pg 141
 - ⁴ Nikos Mourtos, Emily Allen, Introducing Cooperative Learning through a Faculty Instructional Development Program, (JEE October 2001), pg 669
 - ⁵ Dave Ellis, Becoming a Master Student, 13th Edition, (Boston, Wadsworth/Cengage Learning, 2011), pg 231
 - ⁶ Felder, R., "Who Needs These Headaches?", (Success 101, No.4, Fall, 1997)
 - ⁷ George Catalano, Karen Catalano, Transformation: From Teacher-Centered to Student-Centered Engineering Education, (JEE January 1999), pg 60
 - ⁸ Krzysztof Fidkowski, On Oral Exams in Undergraduate Engineering Classes; Fourth Annual Research and Scholarship in Engineering Education Poster Session, October 2009
 - ⁹ Cengiz Gulek, Preparing for High-Stakes Testing; Theory Into Practice, Vol. 42, Number 1, Winter 2003, (Ohio State University College of Education, 2003), pg 43
 - ¹⁰ Steve Roney, Donald Woods, Ideas to Minimize Exam Anxiety, (JEE July 2003), pg 249
 - ¹¹ David Burns, Will I Do as Well on the Final Exam as I Expect? An Examination of Students' Expectations, (Journal of the Scholarship of Teaching and Learning, Vol. 8, No. 3, October 2008), pg. 2

Appendix A

Oral Exam Rubric

| WEIGHT | CRITERIA | SCORE | TOTAL |
|--------------------|---|-------|-------|
| 10% | Assumptions (listed, applicable, correct, understood) | | |
| 20% | Approach (clear understanding of problem, correct method or use of equations and application of assumptions) | | |
| 15% | Equations (correct governing equations listed and used) | | |
| 25% | Variable Treatment (clear understanding of variables, correct values substituted in) | | |
| 5% | Units (correct units in answer, consistent treatment of units) | | |
| 10% | Reasonable Answer (discussion on answer - validity discussed with justification) | | |
| 15% | Algebra/Mathematical Manipulation (all math procedures carried out correctly) | | |
| TOTAL GRADE | | | |

Score for each column is 0-5

0 = no understanding or performance

(had nothing written, no work completed, cannot explain anything, hints won't help at all)

1 = poor understanding or performance (well below average – F)

(had some information written, but full of mistakes or misconceptions, hints don't really help)

2 = fair understanding or performance (below average - D)

(had information written, but several mistakes or misunderstandings, hints help a little, but still unsure)

3 = good understanding or performance (average – C)

(had information written, a few mistakes, but generally OK, hints help a lot, able to get back on track)

4 = high level of understanding or performance (above average – B)

(had information written, minimal mistakes, good understanding, hints barely needed)

5 = excellent or perfect understanding or performance (well above average – A)

(had information written, no mistakes, total understanding, no hints needed)

Appendix B

Post Oral Exam Survey

Please answer with honesty and objectivity. For the first 15 questions (and #17), please circle the feeling you most agree with. There is a chance for free response on the second page.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

1. Compared to conventional testing, the testing procedure used for this first exam facilitates broader learning for me.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

2. Compared to conventional testing, this method of testing facilitates more in-depth learning on the particular concepts that related to my questions.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

3. Compared to conventional testing, this process is much more stressful.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

4. Compared to conventional testing, this method of testing allows for a more accurate reflection of the students' knowledge.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

5. By having to explain my solution in person, I feel this method better prepares me for the professional working environment.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

6. I feel the teaming aspect of this method prevented accurate individual assessment.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

7. I feel the variation in difficulty of the problems made the assessment unfair.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

8. I feel the teaming aspect of this method greatly enhanced my learning.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

9. I feel that completing the oral portion of the exam in a team setting is beneficial.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

10. I feel the process used to form teams was satisfactory.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

11. I feel the team size was appropriate.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

12. The timeframe provided for solving the exam problems was appropriate.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

13. The overall level of difficulty of the exam problems was appropriate.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

14. I feel that completing the oral portion of the exam would be better on an individual basis.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

15. The amount of time provided for the oral exam portion was appropriate.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

16. I have viewed my grade for this exam, online. T or F

If you answered True to #16, please answer #17

17. I feel my grade for this exam is reasonably accurate of my knowledge.

SA = Strongly Agree; A = Agree; N = No opinion/Neutral; D = Disagree; SD = Strong Disagree

18. If you answered D or SD to #17, please explain why:

Free Response to any question on last page.

Appendix C

The take home test helped me learn more instead of just memorizing

The oral exams were very nice because I learned how to apply equations and principles to problems, rather than cramming equations into my head the night before and stumbling on the test.

This style of test was amazing; it reduces stress and increases the amount of valuable material learned during the course.

I liked the way the exams were conducted; it allowed me to reinforce my understanding of subjects by talking to my group. I believe this style of exam was beneficial to all of us because we were able to work together to solve the problems and cross-check out teammates' work.

The group [oral exams] helped me learn the material on a level memorization would never have taken me to.

I liked how [this testing] wasn't just a "memorize the material and apply it to this question in an hour" type of test. It wasn't the cookie cutter stuff that we get in every other class. Instead, harder questions were issued, of which we had to know every little aspect about. Every step, every variable, every assumption, we had to know in order to properly present our findings. Our resources were unlimited, as in the real world, our time constrained, and our knowledge challenged. Excellent method overall.

The oral exams really guaranteed that I got the highest grade possible for an exam as a result of my knowledge, and not what I can remember and write down in an hour. The oral exams really made me have to sit down and learn the material so that I would be able to explain to the professor my methods, answers and how I would adjust my analysis according to changes that could occur in the problem. This most helped me learn the coursework.

The take home exams, although very difficult and time consuming, I feel [helped me] learn a lot more than I would have by just cramming for a regular in class exam. Not only did trying to figure out the problems help, but learning how to do them well enough to be able to explain it to someone really helped get it into my head.

The oral exams were excellent; they forced me to be an expert in the subject matter. After all, once you can teach the material to someone you have a great understanding of it.