

Different Assessment Methods in a Civil Engineering Course

V.Hellum, A.Heimdal, and I.Lande, *University of Agder*

ABSTRACT:

Due to the outbreak of Covid-19, there has been a need for exploring different assessment methods and approaches to teaching. The purpose of the present paper is to share the experiences and discuss three different assessment methods that have been used in the same university course within civil engineering studies over the last three years. The three assessment methods which have been explored are (i) a conventional four-hour unseen exam with calculation tasks performed at the campus, (ii) a 24-hour home exam with calculation tasks and (iii) a portfolio-based assessment with an additional oral exam.

The choice of assessment method clearly influenced the grades where the portfolio-based assessment with the additional oral exam produced more than a grade higher average than the two other assessment methods. Interestingly, there were no indications of gender differences between the assessment methods. Females and males achieved similar average grades in all the introduced modes of assessments.

Based on comparisons between the portfolio results and the following oral exam there are indications that the improved grades were not only a function of an “easier assessment method” but might also be because of better performance which may partly be induced by the mode of assessment.

1 INTRODUCTION

The aim of this paper is to present and discuss the results of a case study on the use of different assessment methods in a course within civil engineering study.

In general, at university level, the students receive marks in most courses based on an assessment method which is described to the students before the start-up of the course. The assessment method, or rather the students' perception of the assessment method is believed to have significant influence on their approaches to learning (Flores, Veiga Simão, Barros, & Pereira, 2015; Struyven, Dochy, & Janssens, 2005). In a discussion held by Gibbs and Simpson (Gibbs & Simpson, 2005) their starting assumption is “there is more leverage to improve teaching though changing aspects of assessment than there is in changing anything else”. For many of us, necessary changes to assessment methods enforced by the covid-19 pandemic could thus be a great opportunity to improve our courses if valid research is used as foundation.

Some assessment methods such as multiple-choice tests or unseen exams are believed by many to promote surface approaches to learning where the students seek “the hidden curriculum”, and student effort is often focused towards figuring out what the lecturer will be looking for during the assessment rather than focusing on understanding the content of the course (Gibbs & Simpson, 2005). Nevertheless, these assessment methods are still used to a great extent especially in courses with large enrolment as these methods often require less resources (Gibbs & Lucas, 1997).

Coursework assessment has in many research programs been shown to result in higher average marks than the unseen exams (Bridges et al., 2002; Gibbs & Lucas, 1997; Woodfield, Earl-Novell, & Solomon, 2005). There are some concerns related to this, for example, is it simply easier to gain a good mark during coursework? And is there extensive cheating during such work? These arguments would suggest less quality embodied in marks from coursework. However, as Gibbs and Lucas argue: “Coursework usually engages students in continuous work during the course”, “Coursework provides students with feedback which helps them to decide what to pay attention to”, “Coursework marks tend to correlate

better with long term measures of learning than do exams”, “Coursework often more closely resembles the kinds of tasks students will undertake after they have graduated than do exams which involve speed work from memory quite unlike most professional practice”. Hence, despite some concerns related to the quality of coursework exams, there are many arguments supporting improved learning from coursework assessment versus unseen exams.

With the reported higher marks from coursework assessment some has raised questions whether this favour females in courses with an extensive use of coursework. Woodfield et al. (Woodfield et al., 2005) discussed this topic based on statements from other authors that males are usually better at taking risks and capturing the big picture, whereas females are better at systematic work and capturing the details which is better suited for coursework without short time constraints. Hence, when larger portions of coursework assessments are introduced, this favour female students. Woodfield et al. however stated that based on their results females outperformed their male peers independent of the assessment method. And that there was no significant difference based within the assessment methods. Salehi et al. (Salehi et al., 2019) did research on gender differences in STEM courses and found that there was a gender gap in favour of males from high stakes unseen exams in undergraduate STEM courses. Their investigations found that a main contributor to this was negative influence by test anxiety which was reported to be more common with females than males. However, the gap decreased when accounting for incoming preparation. For other forms of assessment, including for example coursework, the females outperformed their male counterparts which corroborates the results of Woodfield et al.

Within the engineering study programmes at our university, what can be referred to as the conventional assessment method, is a four-hour school exam with supervision, involving algorithmic problems, which often has a single solution. During this type of exam, the students will face problems which are similar but usually not equal to problems they have faced before. The problems may be closely related to problems which may be expected in industrial settings during professional work after graduation. The problems can be solved by using the same formulas, mathematical/physical/etc. laws or concepts as they have used to solve similar problems during the courses.

With the Covid-19 outbreak and the implementation of social distancing rules the 4-hour school exam was not an option. Hence, different assessment methods were explored. During the course explored herein, the alternative assessment method which was introduced during the first year of the pandemic was a 24-hour home exam. During the second year of the pandemic a portfolio-based assessment followed by an oral examination online was introduced.

The focus of the present paper is merely to present the empirical data and discuss the differences between the deployed assessment methods, how the results differed between the different assessment methods and what might be the possible reasons for this. We believe that sharing all such experiences may help promote discussion and raise the quality of our education. From this the following questions will be asked:

- **How did the different assessment methods influence the marks?**
- **Do the marks from the different assessment methods indicate any gender differences?**
- **Is it possible to use this data to recommend one of the assessment methods over the other?**

2 BACKGROUND

The course used for the case study in the present paper is for first-year students in Civil Engineering during the second semester. The course is on mechanics and strength of materials which involves applied physics and calculation techniques for understanding the behaviour of and assessing structural members subject to different types of loading. It is lectured at a Norwegian University with Norwegian as the language. The time spent in the classroom is divided such that 50% of the time is used for lecturing where theoretical concepts are derived and practical examples are shown, whereas 50% is used for problem solving (exercises) by the students with guidance from the lecturer and teaching assistants. In table 1 some key numbers for the course are displayed.

Table 1 Information about the different courses from 2019, 2020 and 2021

	BYG110	BYG128	BYG128
Year	2019	2020	2021
ECTS	10	7.5	7.5
Total number of students	99	65	88
Number of female students	31	23	28
Number of male students	68	42	60
Lectures [hours pr week]	4	3	3
Excercise [hours pr week]	4	3	3
Number of teaching assistants	4	4	4
Hand-ins (mandatory)	13 (10)	10 (7)	N/A
Hand-ins (graded)	N/A	N/A	10 (8)
Assessment method	4-hour conventional exam	24-hour home exam	Portfolio and oral exam
	N/A		

The course underwent redesign between the first and the second year where it went from 10 ECTS (BYG110) to 7.5 ECTS (BYG128). Some content was thus removed. The same lecturer had all three courses. The number of teacher assistants helping in the exercise hours was the same for all three years. Following completion of the assessment in the course the students are graded from A to F based on their performance. The letter grade is based on the percentage of points achieved: F=0-39%, E=40-49%, D=50-59%, C=60-79%, B=80-89%, A=90-100%. The grading was performed by two persons each year, the lecturer, which was the same all the three years, and three different external examiners, one for each year. Further, a more detailed description of the assessment methods will be given.

3 METHODS

The three different assessment methods used in this case study are:

1. Conventional exam
2. 24-hour home exam
3. Portfolio and oral exam

3.1 Conventional exam

What is here referred to as the conventional exam is a 4-hour exam performed at the university campus with supervision. The students receive the problems to be solved at the startup of the exam period and have 4 hours to complete all the tasks without any help from any other person. The problems involved the same concepts and fundamentals as those solved during exercises in the course. The only allowed aid was a book of formulas. This may be viewed as a high stakes exam since the grade received on this exam will be included in the students' final diploma. In addition to the graded exam, the students had 13 hand-ins, which is approximately 1 hand-in per week, with a requirement of 10 accepted hand-ins to be approved for the exam. The acceptance criteria for the hand-ins were a score of 40%.

3.2 24-hour home exam

During the first year of the pandemic, a 24-hour home exam was introduced. The background for choosing a 24-hour exam was that many students had kids, which were now at home during the day because of the shutdown of schools and kindergartens. Hence, to ensure that the students had sufficient time to complete the exam, a full 24-hours was made available. The number of tasks to be completed was then larger than what was given on the 4-hour exam. A folder including the exam problems were made available to the students at the start-up of the exam, and a folder for submitting the finished exam

was open the next 24 hours where the students should submit their pdf file. To try and mitigate possible collaboration between students, all problems included in the exam were parameterized such that the students were seeking different answers to the problems. In addition, clear instructions were given regarding that collaboration was not allowed. As for the course with the conventional exam, this course also had mandatory hand-ins where 7 out of 10 hand-ins should be accepted to be eligible for the exam. The acceptance criteria were the same as for the previous year.

3.3 Portfolio and oral exam

The second year of the pandemic a portfolio-based assessment followed by an oral exam was introduced. The students would now receive assignments which should be solved and submitted approximately every 1.5 week during the course. Each of the assignments had their own deadlines. These assignments could be solved during the exercises performed with guidance from the lecturer and the teacher assistants. Further, the students could now collaborate and use all available resources. The tasks included in these assignments were similar to those given in previous exams and in most cases relevant for industrial work. Each of the assignments were graded successively and at the end of the course a portfolio of 8 assignments would be selected for the portfolio grading. The average grade of these 8 assignments would then give the portfolio grade which would be the foundation for the final grade. At the end of the course an oral exam was executed. The final grade would then be determined based on the performance on the oral exam, where the grade from the portfolio could be adjusted by 1 grade. Hence, if the portfolio grade was B, the final grade would be A, B or C depending on the performance on the oral exam. Despite only an adjustment could be made following the oral exam, an unofficial grading of the oral exams was performed. This was purely for research purposes to be able to compare the isolated oral exam performance towards the portfolio grade and the final grade. During the oral exam the students were presented with problems they would be given on conventional exams and which they also worked with during the portfolio assignments. They were then asked to explain how they would proceed to solve these problems. Hence, compared to a conventional exam, where a lot of calculation tasks are performed it was not any risk of performing any calculation errors. The purpose of this was to determine their understanding of the material they had delivered during the portfolio assessment. The oral exam was performed as video conference calls where the student and the two examiners joined from different locations and had a duration of approximately 15 minutes.

4 RESULTS AND DISCUSSION

In this section the empirical data, which are based on the achieved grades by the students are presented and discussed. During all three courses some students were enrolled at the start of the semester, but quit the course before the assessment, I.E before attending the final exam or before the hand in of the portfolio. These students are not included in the data.

4.1 Results from different assessment methods

Figure 1 shows the distribution of the final grades obtained from the different assessment methods. From the graph it is seen that the portfolio/oral exam assessment produced a much higher percentage of the two highest grades compared to the two other assessment methods, which is in line with findings in the literature that coursework produces higher grades.

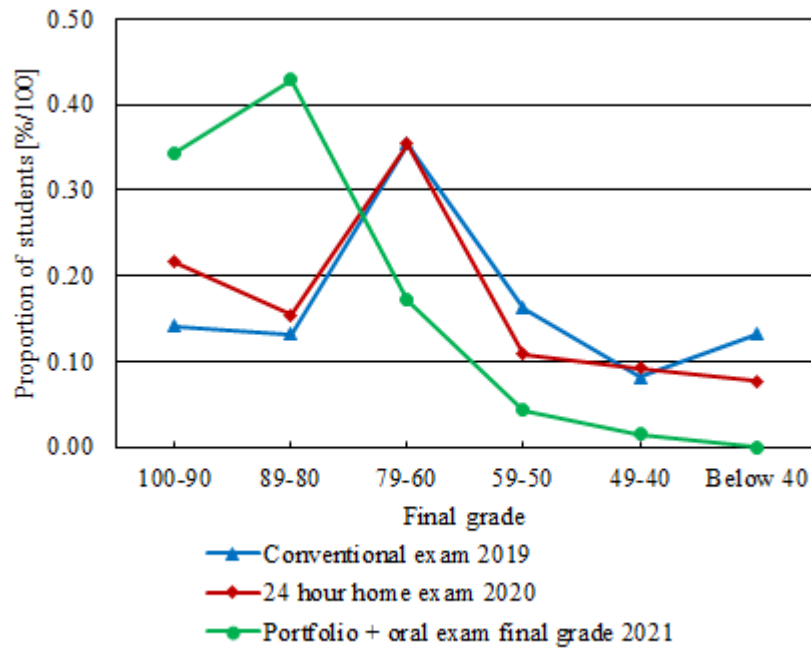


Figure 1 The distribution of the final grades for the three assessment methods

This is also seen when looking at the average grade as shown in Table 2. Here, A equals to 6, B equals to 5, C equals to 4, D equals to 3, E equals to 2 and F equals to 1. Hence the conventional 4-hour exam gave an average grade slightly below C, the 24-hour home exam gives an average grade slightly above C and the final grade from the portfolio + oral exam gave an average grade slightly above B.

Table 2 Average grades from the different assessment methods

Assessment method	Conventional 4-hour exam	24-hour home exam	Final grade Portfolio + Oral exam
Average grade	3.70	4.06	5.04

Hence, the assessment involving coursework and oral exam gave more than 1 average grade better results than both the 24-hour and the 4-hour exams, both of which could be seen as high stakes exams. The grades from the 24-hour home exam were slightly better than the conventional 4-hour exam which might be attributed to the longer time period for executing the exam. Salehi et al. (Salehi et al., 2019) notes “extending the time allowed to complete exams” as a possibility for lowering the test anxiety. On the other hand, the increase in average results could also be attributed to the possibility of collaborating with others during this home exam. Even though the exams were personalized for each student, there were no supervision, so the possibility of interacting with other students or even external helpers were obviously present.

When separating the portfolio grade, the unofficial grade from the oral exam and the final grade from 2021 and comparing this to the results from the conventional exam and the 24-hour home exam some interesting features may be seen as displayed in Figure 2. When looking at the portfolio grade, 70% of the students achieved the best possible results. This is hardly surprising since the students could collaborate and receive guidance while working with the material. And being a course of mechanics and

strength of materials, the problems usually had exact answers. The element of having a bad day on the exam due to for instance test anxiety is not applicable, hence hard work would in most cases be directly translated into a good grade. However, one might also achieve a good grade by working with other high performing students who could help performing the tasks such that in reality one would have a good grade, but limited knowledge. The latter would probably lead to a poor performance on the oral exam. When comparing the isolated oral exam grades towards the conventional exam and the 24-hour home exam, the oral exam produced the highest percentage of the two highest grades of the three assessment methods. On the other hand, the oral exam produced a similar amount of “below 40” or F as the 24-hour home exam. A possible explanation here is as mentioned earlier the effects of collaboration, where some students would benefit from helping others where this would increase their understanding of the material, while some students which are dependent on receiving help have managed to deliver assignments with good results without capturing the essence of the material. The oral exam might not be viewed as a high stakes exam since it would only adjust the portfolio grade by one grade. But test anxiety might still be an issue when you are expected to explain to two lecturers in a short amount of time how you would solve different problems.

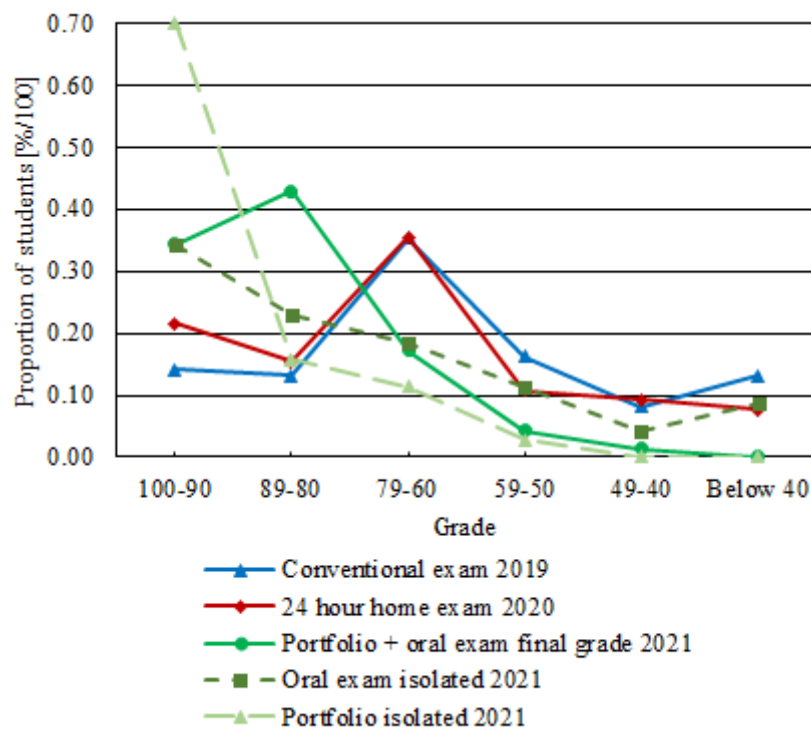


Figure 2 Comparison between the different exam forms

When looking at the average grades also including the isolated portfolio and oral exam grades ref. Table 3, interestingly the average grade from the oral exam significantly exceeds the average grades from the conventional 4-hour exam and the 24-hour home exam. This could indicate that the students did not only perform better because of the portfolio assessment being easier.

Table 3 Average grades including separate portfolio and oral exam grades

Assessment method	Conventional 4-hour exam	24-hour home exam	Final grade Portfolio + Oral exam	Portfolio	Oral exam
Average grade	3.70	4.06	5.04	5.53	4.46

In terms of collaboration during the portfolio assessment, this might be seen as a positive feature since this is often resembling a professional work method. However, as previously stated there might be some students having significantly overperformed on the portfolio due to help from others. The next two graphs are included to address this. Figure 3 shows the difference in grade between the final grade and what was achieved on the oral exam. Bear in mind that the portfolio grade could only be adjusted with one grade. The graph shows that 70% of the students achieved the same final grade as the performance on the oral exam would give. However, 11% of the students ended up with a final grade which was 1 grade better than what they would achieve from the oral exam. Further, 11%, 3% and 4% of the students ended up with a final grade which was 2,3 and 4 grades better than what the performance on the oral exam would give. This means that 4% of the students ended up with a final grade of B while only achieving an F on the oral exam. This might be seen as a flaw in the assessment method used here. I.E., one could argue that there should at least have been a requirement that both the portfolio and the oral exam should be passed. It is reason to believe that the students with the largest gaps between the oral exam grade and the final grade have relied on others to be able to complete the portfolio and have passed the course, with a good grade, with limited knowledge of the content of the course. On the other hand, 70% of the students achieving the same final grade as the achieved grade on the oral exam is very positive. As mentioned earlier, the portfolio/oral exam produced the highest percentage of the two highest grades. As also mentioned, the portfolio mitigates the possibility of having a bad exam day. But this cannot be said about the oral exam. Assuming some of the students had a bad day on the oral exam, it can be viewed as a very good result that 92% of the students are within a gap of 2 grades between the final grade and the grade achieved on the oral exam.

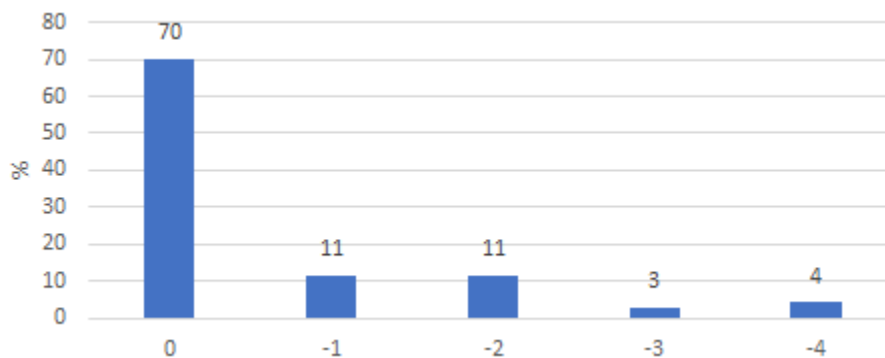


Figure 3 Difference in grade between final grade and oral exam. Negative difference indicates that the grade has gone down after oral exam.

Figure 4 shows the differences in grades between the portfolio and the oral exam. Here it is shown that 4% of the students performed better on the oral exam than what was achieved on the portfolio.

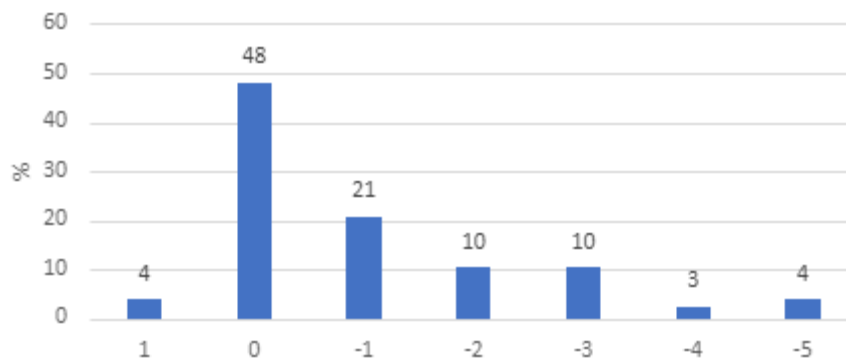


Figure 4 Difference between oral exam and portfolio grade. Negative difference indicates that the grade on the oral exam was lower than the portfolio grade.

To summarize, the assessment method including the portfolio produced markedly higher grades than the two different high stakes exams. The oral exam following the portfolio also produced a higher average grade than the two high stakes exams. The latter may be viewed as a good result in terms of students performing better and not only receiving better grades. There were indications that some students had been given too good grades compared to their performance. This could probably be mitigated by requiring both assessments in the combined portfolio and oral examination to be passed.

4.2 Gender

When comparing the average grades achieved by females (F) and males (M) displayed in Table 4 for the different assessment methods we would argue that there are no indications of any gender gaps from any of the assessment methods in this course.

Table 4 Average results from the different exams for males and females

Assessment method	Conventional 4-hour exam		24-hour home exam		Final grade Portfolio + oral exam		Portfolio		Oral exam	
	F	M	F	M	F	M	F	M	F	M
Average grade	3.61	3.76	4.13	4.02	4.96	5.09	5.52	5.53	4.48	4.44

4.3 Notes on the results and some discussions on the practical aspects

It is emphasized that the results provided represent a small sample size. No information regarding incoming preparation is evaluated and statistics regarding students work effort during the course is not available. There is also no solid knowledge about possible cheating during any of the completed assessment methods. These are all aspects which may have influenced the results. However, the conclusions presented in the following section are mostly consistent with expected outcomes based on available literature.

There was a requirement for students to deliver mandatory hand-ins to be eligible for the 4-hour exam and the 24-hour home exam. Such mandatory hand-ins with pass fail could also be viewed as assessment. However, the big difference is that the threshold for achieving a pass on the mandatory hand-ins were low and the hand-ins were followed by high stakes exams which could draw the students' attention towards the exam.

The portfolio assessment and the oral exam require more resources from the lecturer than the 24-hour exam or the 4-hour exam. For the portfolio assessment, new assignments including solutions had to be produced every one-and-a-half week followed by censoring all the assignments. However, if considering the mandatory hand-ins which was part of the course when the other assessment methods were used, the increase was not significant. For the mandatory hand-ins there were some elements of re-use which was not possible during the portfolio assessment. 3 days were spent on the oral examination where the censoring was done successively. This is comparable to the amount of time spent on censoring the 4-hour exam and the 24-hour exam. It should be noted that the time spent during the different assessment methods were not recorded so the argumentation here is based on estimates from the lecturer.

5 CONCLUSIONS

The investigations of the three different assessment methods leads to the following conclusions being drawn:

- The choice of assessment method clearly influenced the grades where the portfolio-based assessment with the additional oral exam produced more than a grade higher average than the two other assessment methods.
- There were no indications of gender differences between the assessment methods. Females and males achieved similar average grades in all the introduced modes of assessments.
- Based on comparisons between the portfolio results and the following oral exam there are indications that the improved grades were not only a function of an “easier assessment method” but might also be because of better performance which may partly be induced by the mode of assessment.

For future research, the following aspects could be investigated:

- The influence of student related parameters, such as age, living conditions and scientific level.
- Similar investigations in other courses.
- Include qualitative data from the students on their perception on different assessment methods.

REFERENCES

- Bridges, P., Cooper, A., Evanson, P., Haines, C., Jenkins, D., Scurry, D., ... Yorke, M. (2002). Coursework Marks High, Examination Marks Low: Discuss. *Assessment & Evaluation in Higher Education*, 27(1), 35–48. <https://doi.org/10.1080/02602930120105045>
- Flores, M. A., Veiga Simão, A. M., Barros, A., & Pereira, D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: a study in higher education. *Studies in Higher Education*, 40(9), 1523–1534. <https://doi.org/10.1080/03075079.2014.881348>
- Gibbs, G., & Lucas, L. (1997). Coursework Assessment, Class Size and Student Performance: 1984-94. *Journal of Further and Higher Education*, 21(2), 183–192. <https://doi.org/10.1080/0309877970210204>
- Gibbs, G., & Simpson, C. (2005). Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education*, (1), 3–31. Retrieved from <https://eprints.glos.ac.uk/3609/>
- Salehi, S., Cotner, S., Azarin, S. M., Carlson, E. E., Driessen, M., Ferry, V. E., ... Ballen, C. J. (2019). Gender Performance Gaps Across Different Assessment Methods and the Underlying Mechanisms: The Case of Incoming Preparation and Test Anxiety . *Frontiers in Education* , Vol. 4. Retrieved from <https://www.frontiersin.org/articles/10.3389/feduc.2019.00107>
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education*, 30(4), 325–341. <https://doi.org/10.1080/02602930500099102>
- Woodfield, R., Earl-Novell, S., & Solomon, L. (2005). Gender and mode of assessment at university: should we assume female students are better suited to coursework and males to unseen examinations? *Assessment & Evaluation in Higher Education*, 30(1), 35–50. <https://doi.org/10.1080/0260293042003243887>