The relationship between ICT based formative assessment and academic achievement in a Mechanics of Materials course

M.A.H.A. Meijers[1]
Avans University of Applied Sciences
Breda, The Netherlands

P.P.J.L. Verkoeijen
Avans University of Applied Sciences
Breda, The Netherlands

Conference Key Areas: Open and Online Teaching and Learning, Integrated learning environments for the digital native learners

Keywords: formative assessment, feedback, personal, ICT, student achievement

ABSTRACT

Formative assessment can be understood as assessment for learning rather than of learning. Research has demonstrated that formative assessment/feedback is strongly and causally related to students' achievement in higher education. At Avans University of Applied Sciences, the first-year course Mechanics of Materials 2 (MoM2) does not contain formative assessment due to constrained personnel and financial resources. In this study, we introduced formative assessment in the MoM2 course using an innovative ICT tool. This tool provides students with personalized feedback on both task and process level to improve their learning and to enhance adaptive teacher class instruction.

To assess the impact of the ICT based formative assessment, we obtained the final course exam pass rate, the average grade and student satisfaction on the 2018-2019 course MoM2. We compared these outcomes to those in the previous 3 runs of this course (2017-2018, 2016-2017, 2015-2016). These courses were identical to the most recent version of MoM2 but they did not include ICT based formative assessment. In addition, we compared the 2018-2019 run of MoM2 with the 2018-2019 run of Mechanics of Materials 1 (MoM1), which is a highly similar course, which does contain a form of non-ICT based formative assessment.

The results show that the use of ICT based formative assessment did not improve academic achievement in the course MoM2 2018/2019 compared with previous runs of the same course and did not improve academic achievement compared to the similar course MoM1 2018/2019.

1
M.A.H.A. Meijers
maha.meijers@avans.nl

1 INTRODUCTION

Summative assessment usually comes at the end of a course to evaluate student performance with respect to learning goals. In contrast, formative assessment takes place during a course with the aim to provide feedback that students help to observe, enhance and accelerate their learning process (e.g., [1] and [2]). In addition teachers, use formative assessment to adapt their instruction and to enhance student learning.
There are several reasons to assume that meaningful learning, self-efficacy and self-regulated learning are supported by formative assessment. First, the provided feedback informs the student about the learning goals and the criteria used during summative assessment. In addition, the formative assessment also informs the student about the current performance level with respect to those learning goals. This helps students to direct and/or adapt their future learning effort and learning strategies towards the learning goals [1]. Research has demonstrated that formative assessment/feedback is strongly and causally related to students' achievement in higher education. For example, review articles by Black [3], Hattie [4] and Schneider and Preckel [5] suggest that formative assessment may have a strong positive effect on academic achievement. In [4] and [5] feedback – which is a crucial part of formative assessment – is even rated as one of the top ten factors related to academic achievement.

At Avans University of Applied Sciences Mechatronics (BSc.) program, the first-year course Mechanics of Materials 2 (MoM2) (2 ECTS) contains limited formative assessment due to constrained personnel and financial resources. Students only receive task level feedback: they receive worked-out solutions to their homework one week after submission. Additionally there is the possibility for students to review the results of the summative end-of-course exam. Feedback at the process level is absent. In this study, we introduced formative assessment in the MoM2 course using an innovative ICT platform. This platform provides students with personalized feedback on both task and process level to improve their learning and to enhance adaptive teacher class instruction, while limiting the necessary financial and personnel resources. Our research question was if this ICT based feedback improves academic achievement in the course MoM2.

2 METHODS

2.1 Course and participants

At Avans University of Applied Sciences, the first-year of the Mechatronics curriculum is divided into 4 quarters, each of which consisting of 10 weeks. In these quarters, the first nine weeks are reserved for educational activities and end-of-course exams; the final week is reserved for re-sit exams. MoM is covered in two courses during the first-year: MoM1 is conducted in the first quarter and MoM2 in the third quarter. Both courses have a work load of 2 ECTS and within their quarter each MoM course runs parallel to other courses and projects with a total course load of 13 ECTS. Hence, each MoM course covers only a small part of the load in each quarter.

Furthermore, MoM2 is conducted after a first informal selection point in Dutch higher education, i.e., February 1st. Students who have not been able to meet a program's standard, have the opportunity to leave the program without financial consequences. This means that MoM2 is likely to contain a subset of students who performed relatively well.

In MoM2, as well as in MoM1, classes with a maximum of 32 students take part in 7 instructional sessions (1 per week) of 90 minutes. During these sessions, new theory and problem solving strategies are explained and students work on practice problems. The MoM2 course in 2018-2019 was similar in content and design [6] as the MoM2 course in previous years. Furthermore, the lecturer of MoM1 and MoM2 was the same in all courses considered in this study.

Participants in this study were first year undergraduate students that actively participated for the first time in MoM1 and MoM2. Active participation implies that a student made his/her homework for at least one course meeting. Hence, non-active students are those who did not make any homework assignment at all during a course. Students were allowed to opt-out from this study and provide their homework according to the procedures of previous years.

2.2 Intervention

The intervention consisted of formative feedback based on data from our ICT tool [7]. Each week, this ICT tool assigned a unique set of practice problems to each student. Each set consisted of 3 to 6 problems. The ICT tool provided feedback to students in the form of either adaptive explanatory or socrative feedback. The explanatory feedback was to some degree dependent to the answer given to problems provided by the tool. For example, the tool can recognize when a "sine" is used instead of a "cosine" and will provide feedback to the student to review the trigonometric relations used in solving the problem. Socrative feedback was provided when students asked for a hint in the ICT tool. A hint consists of a decomposition of the stated problem according to the problem solving procedure. The student provided answers to partial problems and received explanatory feedback.

Additionally students received a grade (1 to 10) for each week's set of practice problems to give them insight into their progress during the course.

Our ICT tool [7] provides feedback to the lecturer at both group and student level. Learning performance and time-on-task were used during the course to perform two weekly interventions. Those students who spent less than 2 hours on their homework assignments and performed below 40% on the homework assignment were separately addressed after the instruction in the following week. These student have been asked if they were aware of their underperformance and if they encountered any problems during their homework. Based on the average performance of the students on the problems in the homework assignments, the problem with the lowest performance was selected by the lecturer. This problem was then discussed in class by the lecture in the successive weekly instruction.

2.3 Design

To assess the impact of the ICT based formative assessment, indicators of student achievement and student satisfaction have been obtained from the registrar for the course MoM2 for the academic years 2015-2016, 2016-2017, 2017-2018 and 2018-2019. We compared these outcomes to those in the previous 3 runs of this course (2017-2018, 2016-2017, 2015-2016). These courses were identical to the most recent version of MoM 2 but they did not include ICT based formative assessment. In addition, we compared the 2018-2019 run of MoM 2 with the 2018-2019 run of MoM 1. Both courses are comparable in terms of content [6], types of homework assignment, lecturer and difficulty level and format of the final end-of-course exam. Both courses only differ in the way formative feedback is administered. This is provided by the teacher without assistance of the ICT tool in MoM1 and with the ICT tool [7] in MoM2. Hence, this within-cohort comparison allows for an additional assessment of the effectiveness of our ICT tool apart from the cross-cohort comparison in MoM2.which is an highly similar course, which does contain a form of non-ICT based formative assessment.

Student achievement has been measured with pass rates and the grade on the final end-of-course exam. The grade is expressed as a figure between1 and 10 (note that a grade of at least 5.5 corresponds to a pass). Pass rates and grades of re-sit exams have not been taken into account. Furthermore, any assigned compensation (performed homework) to the final end-of-course exam grade has not been taken into account when determining the pass rate or grade.

With respect to student achievement, we hypothesized the following:

1. The implementation of ICT based formative feedback in the 2018-2019 course MoM2 will enhance students' achievement relative to the MoM2 course in the academic years 2015-2016, 2016-2017 and 2017-2018, where formative feedback is hardly present. I Specifically, for the final end-of-course exam of sterkteleer 2, we expect the following pattern for pass rate (PR), mean grade (M) and median grade (Med):

$$\begin{aligned} \mathrm{PR-M-Med\ (2018-2019)} > \ & [\mathrm{PR-M-Med\ (2015-2016)} \\ = \ & \mathrm{PR-M-Med\ (2016-2017)} \\ = \ & \mathrm{PR-M-Med\ (2017-2018)}] \end{aligned}$$

2. H2: The implementation of ICT based formative feedback in the 2018-2019 course MoM2 will lead to a comparable or better student achievement with respect to the 2018-2019 course MoM1, where formative feedback is provided by lecturers instead of an ICT tool. Specifically, for the final end-of-course exam, we expect the following pattern for pass rate (PR), mean grade (M) and median grade (Med):

$$\text{PR} - \text{M} - \text{Med (Sterkteleer 2 2018} - \text{2019)} \geq$$
$$\text{PR} - \text{M} - \text{Med (Sterkteleer 1 2018} - \text{2019)]}$$

For explanatory purposes student satisfaction and perceived competence has been measured and analytical data from our ICT tool is collected. Student satisfaction has been measured by means of a standardized questionnaire that is similar over the academic years examined in this study. In this questionnaire participants responded to various statements and in the present study we will focus on the following two:

1. “The teaching aids (blackboard, books, lecture notes, ICT applications, etc.) during the project/subject /practical exercises supported my studies”
2. “The methods used during the project/subject/practical exercises contributed to my learning.”

Participants provided an answer to each of these statements on a 5 point-Likert scale ranging from 1= “completely disagree” to 5 = “completely agree”. Per student, the mean score on these two variables has been used as a proxy of student satisfaction with respect to the intervention.

In the 2018-2019 run of MoM2, students’ perceived competence has been measured halfway (week 4 of 7) through the experiment and at the end (week 7 of 7) of the experiment, a week before the final end-of-course exam. This allowed us to explore the development of perceived competence throughout the course. Perceived competence was measured with a short 3-statement questionnaire:

1. “I am not confident that I will be able to master the content that is covered by the final end-of-course exam of Mechanics of Materials 2”
2. “I am confident that I will achieve my learning goals for the course Mechanics of Materials 2”
3. “I am confident that I will pass the final end-of-course exam of Mechanics of Materials 2”

Participants provided an answer to each of these statements on a 5 point-Likert scale ranging from 1= “completely disagree” to 5 = “completely agree”. Per student, the mean score on these three variables has been used as a proxy of perceived competence.

Finally analytical data (i.e., homework duration and performance), from our ICT has been collected for the 2018-2019 run of MoM2.

The design of this study has been pre-registered [8] to exclude confirmation bias and the negative effects of researchers’ degrees of freedom.

3 RESULTS

3.1 Descriptive statistics

Descriptive statistics of student achievement, student perceived competence and student satisfaction in MoM1 and MoM2 can be found in Table 1, Table 2 and Table 3 respectively.

Table 1. Descriptive statistics of student achievement

Course	MoM2				MoM1
Year	2015 2016	2016 2017	2017 2018	2018 2019	2018 2019
Proportion of active students	NA	NA	NA	100%	NA
Pass rates	0.30	0.71	0.65	0.63	0.76
End-of-course grade mean	4.0	6.2	5.8	5.7	6.7
End-of-course grade median	3.8	6.5	6.3	5.9	7.3
End-of-course grade SD	2.1	2.3	2.1	1.9	2.1
End-of-course grade minimum	1	1	1	1	1
End-of-course grade maximum	8.8	9.6	9.7	9.7	10

Table 2. Descriptive statistics of student perceived competence (MoM2 2018).

	Week 4	Week 7
n (number of students)	53	35
Perceived competence mean	2.8	2.9
Perceived competence median	2.7	3.0
Perceived competence SD	.5	.5
Perceived competence minimum	1.3	2.0
Perceived competence maximum	4.0	4.6

Table 3. Descriptive statistics of student satisfaction

Year 2018/2019	Teaching aids	Teaching methods
n	42	42
Mean	3.4	3.7
SD	1.0	0.8

Lastly, for the MoM2 course in the academic year 2018-2019, the relevant descriptive statistics of the two variables related to homework performance (i.e., homework duration and performance) are provided in Table 4.

Table 4. Descriptive statistics of student homework.

	Duration (min)	Performance
n (number of students)	52	52
Mean	67	4.1
Median	76	3.8
SD	17	2.0
minimum	1.5	0.7
maximum	135	8.8

3.2 Confirmatory analyses

To test **hypothesis 1** (see section 2.3), we carried out the analyses according to the pre-registration [8]. A logistic regression analysis on the pas-fail scores, showed a significant relationship between Course year and the probability of passing the test, χ^2 (3) = 30.329, p < .05, Nagelkerke R-square = .168. To follow up this significant effect, we performed a repeated contrast analysis. This analyses showed that the probability of passing the test was significantly lower in the Course year 2015-2016 than in the Course year 2016-2017, *Wald* χ^2 (1) = 22.165, p < .05, *Exp(B)* =.149 [95% CI EXP(B) = .067; .329]. The other two comparisons of the pass rate, i.e., Course year 2016-2017 with 2017-2018 and Course year 2017-2018 with Course year 2018-2019, did not yield significant outcomes.

Furthermore, we performed a single-factor between-subjects Analysis of Variance (ANOVA) on to compare the course years on the mean grades. The ANOVA revealed a significant relationship between Course year and grade, F(3, 230) = 12.481, p < .05, eta-squared = .14. Bonferroni corrected comparisons showed that the mean grade in Course year 2015-2016 was lower than in the Course year 2016-2017, *Mean difference* = 2.250, *Standard error* = .388, p < .05. The other two comparisons of the mean grades, i.e., Course year 2016-2017 with 2017-2018 and Course year 2017-2018 with Course year 2018-2019, did not yield significant outcomes.

A Kruskal-Wallis test on the median grades per course year revealed a significant effect, $H(3) = 31.633$, $p < .05$. Similar to the outcomes for the pass rates and the mean grades, the median grade was significantly lower in the Course year 2015-2016 than in the Course year 2017-2018, $H(1) = 66.079$, $p < .05$. The other two comparisons did not yield significant outcomes.

Taken together, the outcomes of the analyses on the pass rates, the mean grades and the median grades were highly consistent. Performance was worse in the Course year 2015-2016 than in the Course year 2016-2017 but from 2016-2017 onwards performance remained constant. This was contrary to our hypothesis as we expected to performance to be better in the Course year 2018-2019, when the intervention took place, than in the previous years.

To test **Hypothesis 2** (see section 2.3) we planned to perform three analyses to compare MoM1 and MoM2: a logistic regression on the pass-fail scores, a repeated measures ANOVA on the mean grades, and a Wilcoxon test on the median grades. However, the descriptive statistics in Table 1 actually showed that performance according to these three criteria was actually *better* in MoM1 than in MoM2, the course in which the intervention took place. Because the results are diametrically opposed to our one-tailed hypothesis, we refrained from conducting the planned analyses for hypothesis testing purposes. After all, the obtained p-values would be uninformative with respect to our hypothesis. It should be noted that we carried out the analyses for exploratory reasons and these analyses showed that pass rate, mean grade and median grade were higher in MoM1 than MoM2. However, and consistent with the outcomes of the analyses we performed to assess out first hypothesis, the outcomes for the MoM1 versus MoM2 comparison were inconsistent with our expectation that students would perform better in MoM2 than in MoM1.

3.3 Exploratory analyses

A repeated measures ANOVA on the perceived competence scores, showed that mean perceived competence was comparable at week 4 and week 7, $F(1, 34) = .090$, $p > .05$, partial eta squared = .003.

Lastly, we explored the relationship between the total time spent on homework collapsed across the six meetings of the course and the total grade received for the homework assignments (minimum score = 0; maximum score = 60) with performance on the end-of-course grade of MoM2. A multiple linear regression analysis revealed that time spent on homework and homework grade together had a strong positive relationship with end-of-course exam, $F(2, 49) = 7.877$, $p < .05$, r-square = .24. Both time spent on homework (Standardized beta = .465) and homework grade (Standardized beta = .033) were positively associated with the end-of-course exam grade after controlling for the other variable in the model. However, neither the unique contribution of time spent on homework to the end-of-course exam grade, not the unique contribution of homework grade was statistically significant. This was due to an extremely high positive relationship between time spent on homework and homework grade, r = .856.

4 DISCUSSION

In the present study, we examined whether the introduction of an ICT based formative assessment in a MoM course would be associated with an increase in academic achievement. The lecturer of the course MoM2 was heavily involved in this study and much attention has been paid to the proper implementation of the ICT tool. Therefore, we think it is reasonable to conclude that the “treatment” has been delivered according to plan. Nevertheless, the use of ICT based formative assessment did not improve (but also did not deteriorate) academic achievement in the course MoM2 2018/2019 compared with previous runs of the same course and did not improve academic achievement compared to the similar course MoM1 2018/2019. These findings were not in line with our expectations. For this, we have two explanations.

First, the entire instructional design (i.e., worked examples, ordering of practice problems, introduction of theory) within the MoM2 course were of high quality. The same could be said about the lecturer who is an experienced, knowledgeable and skilful teacher. Furthermore, and probably as a result, pass rates of previous runs of MoM2 were already relatively high (>.65, except for the run 2015/2016). In hindsight, therefore, it may be possible that given the quality of the course and the lecturer, there might have been limited room for our ICT-based formative assessment to enhance the end-of-course performance.

A second explanation might lie in suboptimal aspects of the way in which the formative assessment was delivered. For one, from conversations with students, we learned that they perceived their homework results in part as summative instead of purely formative assessment. This might have hindered them to learn from the mistakes they made on the practice problems. As a result, they might have learned less from the practice problem than if they had taken a more formative perspective. This might explain why the perceived competence did not increase during the course (Table 2).

Second, students indicated that they had some problems with the ICT tool, although this is not directly clear form the student satisfaction results (Table 3). They indicated that they could not study together with peers due to personalization of the practice problems and that they spent a lot of time on troubleshooting when they made mistakes. The later arose because the ICT tool returns correct explanations; it does not always indicate why an erroneous student response is incorrect. The way students perceived the feedback, i.e., partly summative, and technical characteristics of the ICT tool might have stood in the way of realizing the full potential of the ICT-based formative assessment.

The use of the ICT tool will be improved with the results of this study. Improvements will focus (among others) on the use of homework results to learn from errors. A new study on the implementation of the enhanced use of the ICT tool will take place in the 2nd quarter of 2019.

5 ACKNOWLEDGMENTS

This research has been carried out with help of a Comenius Teaching Fellow grant provided by the Netherlands Initiative for Education Research (NRO) part of Netherlands Organisation for Scientific Research (NWO).

REFERENCES

[1] Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, *4*(3), 365-379.

[2] Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, *18*(2), 119-144.

[3] Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *92*(1), 81-90.

[4] Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. routledge.

[5] Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological bulletin*, *143*(6), 565.

[6] Hibbeler, R. C. (2017). *Sterkteleer* (9th ed.). Amsterdam, The Netherlands: Pearson Benelux.

[7] Features for Educators | Mastering Engineering | Pearson. (n.d.-b). Retrieved April 14, 2019, from https://www.pearsonmylabandmastering.com /northamerica/masteringengineering/educators/features/index.html.

[8] Meijers, M., Verkoeijen, P., & Musters, V. (2019, January 22). From mechanics for a class to mechanics for a student. Retrieved from https://osf.io/yz25d