Different types of assessments and their effect on students' learning and workload in remote learning

Ali Abolfathi[1]*, Natasha Tahir[1], Khaled Ahmed[1]

[1] Department of Mechanical Engineering, UCL (University College London), London, UK

Abstract: The Covid-19 pandemic disrupted the normal mode of teaching and a remote learning model was substituted in the academic year of 2020-21. As part of the move, a series of additional summative assessments were introduced in many modules of the host department of this study to ensure students' engagement and provide them with feedback while they were learning remotely. Coursework, online quizzes and online exams were used to assess students' learning in this academic year. A student workload model was used to predict the effect of assessments and ensure students have enough time to study all their modules. Surveys were used to determine how different forms of assessment contributed to students learning as well as their impact on students' workload. Questionnaires were designed to collect information on the number of hours students spent on different forms of assessments, how they perceived the effect of those assessments on their learning and how well they have learned different subjects.

Irrespective of the weighting of the assessments and their types, students have treated all summative assessments very seriously and spent considerable time on even small pieces of assessments such as weekly online quizzes with a very low weighting. They perceived completing harder and longer assessments contributed relatively less to their learning. Amongst all three types of assessment, students reported reviewing for online exams helped them the most to enhance their learning.

Keywords; summative assessment, coursework, workload, learning, students' engagement, quiz.

**Correspondence to: C.O. Dr Ali Abolfathi, Department of Mechanical Engineering, UCL. E-mail: a.abolfathi@ucl.ac.uk*

1. INTRODUCTION

In the academic year of 2020-21, the teaching was delivered completely online in the host department of this research, while students were away. New summative assessments in form of quizzes and coursework were introduced to ensure students' engagement and assess their learning while they were away. A workload model was used by tutors where the time requirements for studying and completing different assessments were taken into consideration to ensure students spent similar time studying different modules and assessments would not overload students. The research aims to understand how these new assessments have affected students. Workload and the time they reported spending on different forms of assessment were surveyed. Furthermore, students were questioned about their perception of how completing those assessments has helped them to learn to evaluate the effectiveness of introduced assessments.

A popular framework, the Assessment For Learning framework (AFL) has well-documented benefits, in which learning is a priority, as opposed to individual capability (William, 2004). The purpose is to understand where they are in their learning. Thus, AFL is about reviewing and acting on assessments whereas Assessment Of Learning (AOL) is about a student's ability (López-Pastor et al., 2011).

It is argued that classroom assessment prevails under formative assessments (FA) (Shepard, 2000) as it follows AFL. It has been shown that performance would improve with the incorporation of FA (López-Pastor et al., 2011). However, there may be limited engagement as students have fewer incentives to partake. Comparatively, summative assessments (SA) evaluate performance, though it can, alongside FA, provide 'meaningful performance goals' (McTighe and O'Connor, 2005). This motivates students to excel and provides a measure on which they can reflect. Formative assessments incorporate feedback to nurture academic growth. Furthermore, it differentiates final versions, whereas FA may sometimes be seen as a draft version of work and not possess the same quality as SA.

The above assessments can be incorporated continuously in the year or concentrated into one period. Studies show students prefer continuous assessments (CA) more than teachers and it is beneficial to introduce such assessments if incorporated properly (Olubukola and Bankole, 2015, Hernández, 2012). Furthermore, CA allows teachers to adapt materials better to suit students and target their weaknesses (Ababio and Dumba, 2013). With CA, greater scores are achieved as feedback is provided and improves students' study habits, specifically, for typically low achievers (Onuka and Onabamiro, 2010, Iqbal et al., 2017).

Although an effective assessment regime may be continuous and involve multiple types of assessment, it must not overload students. The excessive workload would have adverse academic and personal implications. In this study, the effectiveness of summative assessments and their workload on second and third-year students are examined. In the next section, the methodology of the study is introduced. In section 3, results and analysis are provided and the paper finishes with conclusions.

2. METHODOLOGY

During the 2020-21 academic year, undergraduate mechanical engineering students were surveyed to collect data on the effectiveness of assessments and students' workload. Two surveys were sent mid-year to year three, and two during the summer, to year two and three students. The surveys were set up on Google forms. The first two were shared via social media apps and the latter two were emailed to students.

Four surveys were sent out to third-year students during teaching terms which resulted in 343 responses. To understand learning preferences, they asked about preferred choices for assessments. It included questions about the time taken to complete assessments of different modules and to study the subjects. These also had comment sections to elaborate on their opinion about effective modes of teaching. Surveys in the summer collected 46 responses; for year three, the number of responses was 32 and for year two was 14.

Significant anomalies of data points were excluded from the datasets rather than all the data points from that respondent, e.g. if a respondent provided an anomaly for the time to do a quiz, it is excluded from the quiz dataset, but the respondent's data would be included for other assessments. Thus, the data points for different assessments may be different. Data on durations were converted to hours, with a working day assumed to be eight hours, for responses provided in days. Students scored their understanding out of five where an understanding equivalent to a first degree was ranked the highest understanding level with a value of 5 and failing a module ranked the lowest with a value of 1. Questionnaires were collected before grade release, so ratings were based on student perception.

Assessments were divided into four categories: coursework, exam, quiz and others. Projects and written assessments were categorised as coursework. Assessments that required answering short questions were categorised as quizzes. Presentations and miscellaneous assessments were grouped as 'other'. For example, module N assessments consisted of six online quizzes, four scenarios, one written assessment and one self-reflection. Scenarios is a week-long project which is completed in teams and either a written report is submitted, a physical item was produced, or findings were presented. This was categorised as six quizzes and five coursework as scenarios are project-based in which students can confer with each other. Some modules, such as O and P consisted of two sub-modules. For these, students understanding of each sub-module was questioned and averaged so, a decimal is reported for understanding level e.g, 3.5.

To obtain the total time spent on a module, the reported time of different components of that module are combined. Where there are multiple assessments of the same weighting and type, the average time to complete each assessment was surveyed, which was then multiplied by the number of assessments. If the weighting was different, e.g., two coursework submissions with one being 10% and the other 30%, the time taken is considered separately.

For analysis, three factors are considered: time taken to complete an assessment, students' perception of the assessment's contribution to their learning and, how well students believe they understood the module. The analysis helps to draw conclusions on the effectiveness of an assessment or combination of assessments while accounting for the workload. The efficacy of assessments in an AFL framework can be evaluated by how well modules are understood and how well assessments contribute to learning.

3. ANALYSIS AND RESULTS

Year three students were assessed through a mixture of open-book exams, coursework and quizzes. Three third-year core modules that are named O, P and Q are considered for this analysis. All three modules are analytical and are of a similar level of difficulty. When considering the time spent on each assessment of different types, exams, including the preparation time, were reported as the most time-consuming assessment type by year three students, followed by coursework and quizzes, as shown in Figure 1(a). However, when accounting for the total time spent on different types of assessment, students spent most of their time on coursework followed by quizzes and exams as depicted in Figure 1(b). This is because there were more courseworks than exams.

Coursework amounted to 71% in weighting for the modules analysed and only 20% was exams. However, quizzes, which have taken more of their time to complete compared to exams, only counted for 9% of their total marks in core modules. Quizzes were introduced as a means to encourage students to participate in the learning process while they were away and studying online. There is no correlation between the weighting of an assessment type and the level of effort that students devote to it when they strive for perfection. This may provide an effective tool to engage students in the learning process but unintended workload may undermine their contributions to other modules and negatively affect their wellbeing.

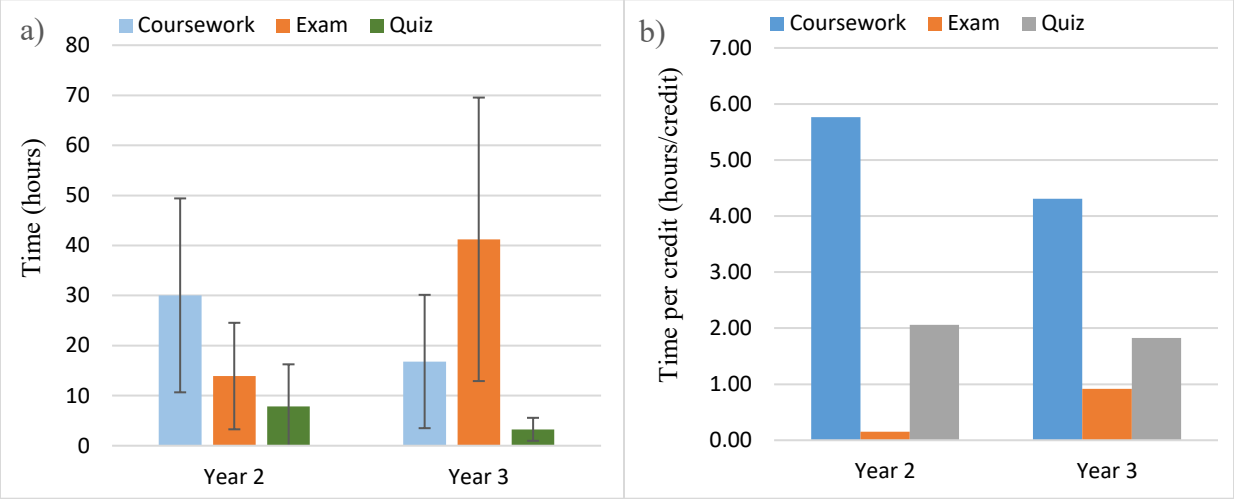

Figure 1. Time spent on different types of assessments for year two and three students. (a) Average time spent on each assessment. (b) Averaged total time spent on each type of assessment normalised by total credit of modules (year 2: six modules with a total credit of 90 are analysed, year 3: three modules with a total credit of 45 are analysed).

Figure 2 shows students' perception of how completing a piece of assessment contributes to their learning as a function of the time they spent on them in each module. Although the scatter of data is very wide, a second-order polynomial fit is used to obtain the trend. Students who spent more time on preparation for exams reported a better understanding of the modules due to the extra time they dedicated. However, the opposite is true for coursework and quizzes, with the contribution to students' learning reduces for those students who took longer for them to complete the assessments. Although students report that the coursework and quizzes have not contributed to their learning in the same way as exams did, it should be noted that coursework and quizzes proceed exams. When students were revising for exams, they had benefited from their studies during the term time to complete quizzes and coursework. Considering the weighting of assessments and the workload student reported, a reduced number of quizzes and shorter coursework may improve students' learning experience in this case.

Students have scored their understanding of all three core third-year modules similarly. The module with the highest understanding in year 3, illustrated in Figure 3, is Q with an average rating of 3.72. The assessment of module Q constituted of 60% exam, 28% coursework and 12% quiz. It is followed by module P with a reported understanding level of 3.66 which was assessed completely by coursework. Students scored their understanding of module O 3.52 which is the

lowest among the three core modules. The module consisted of two parts taught by two academics. There were two weekly summative assessments: an online quiz set by lecturer O-S1 and a short exercise set by Lecturer O-S2. While the quiz was timed and students have to answer one or two questions online on Thursdays, students had a week to complete the exercise question of the weekly task. On average, students scored the contribution to learning 3.32 for quizzes and 3.72 for weekly tasks. Students found weekly online quizzes difficult and stressful and reported spending several hours studying the relevant part of the module to take the quiz. Although one may expect studying for a longer period would improve students learning, it seems that the difficulty level of an assessment would negatively affect students' perceptions of how it has helped them to learn a topic.

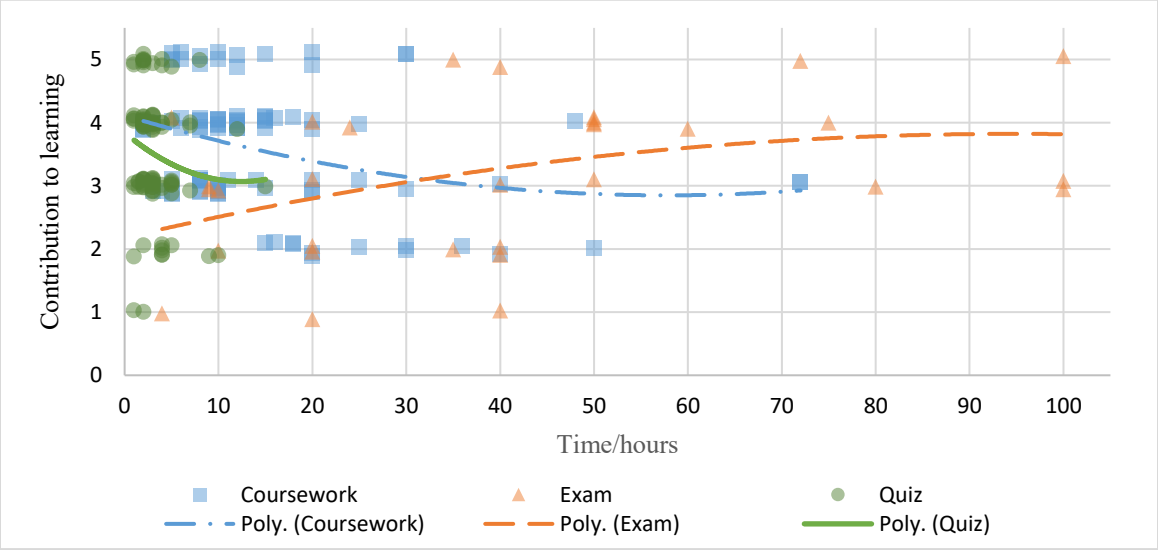

Figure 2: Contribution of different types of assessment to students' learning as a function of the time spent on them for year three students.

The average total time reported by students to complete all assessments for Q, O and P are 114.4, 113.4 and 95.5 hours respectively. In the planning stage, a workload model was used to have a similar workload for different modules that carry the same credits and the time spent on them falls within the prescribed limits. This was done by module tutors predicting the time required to complete different assessments and study the teaching material. While students reported spending a similar number of hours on modules Q and O, the levels of understanding reported were different. However, the average marks attained were similar for these two modules, 72% for Q and 70% for O. Both modules are generally of the same level of difficulty but students complain about the difficulty level of the weekly quizzes of module O in particular. Furthermore, in module O students have a weekly quiz but in module Q they only have two quizzes. It appears that the difficulty and the number of quizzes in form of summative assessments have hindered students' learning experience and while they achieved similar marks overall for both modules, they reported a lower level of understanding when there are multiple small summative assessments. Furthermore, students spent considerably more time on online quizzes than what was predicted by the tutor in the workload model.

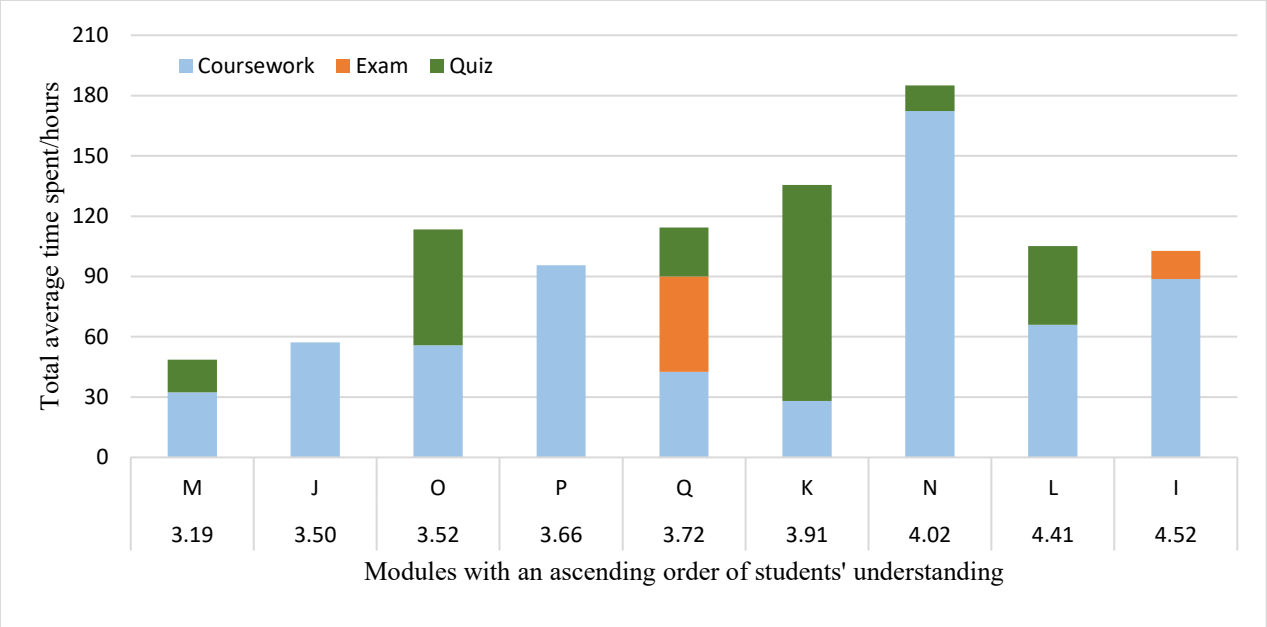

Figure 3: Total average time spent on different modules as a function of the students' perceptions of understanding for year two and three.

Year two modules that are investigated in this study are named I, J, K, L, M, and N. All modules are analytical except module N which is on engineering design. Year 2 students spent most of their time per module on coursework (30 hours), followed by exams (14 hours) and quizzes (8 hours) per piece of assessment on average. Each student on average spent 519 hours on all coursework, 185 hours on quizzes and 14 hours on an exam. This is in line with the fact that 89.2% of total weightings for all modules were attributed to coursework, 7.5% for quizzes and 3.3% for exams, however, relatively more time was spent on quizzes than coursework considering their respective weighting. A gradual increase in average time spent on each module as the understanding increases is noticeable in Figure 3 with an exception for modules L and I which with a similar average time, students reported the highest level of understanding of 4.41 and 4.52. Coursework was the main item of assessment that students' time was dedicated to. While module L included a series of quizzes, module I had a final exam for which students spent relatively a shorter period of time revising. Comparing modules K and N, students reported a better understanding where the coursework was the dominant mode of assessment rather than quizzes. Overall, students reported a better understanding when completing all assessments for a module has taken about 100 hours for them and comprised of two modes of assessment, either a quiz and coursework or exam and coursework while the latter provided a better means to learn a subject.

Students' perception of how different means of assessment has contributed to their learning is shown in Figure 4 for year two students. Each data point represents an entry for an assessment related to any modules that have been taught in year two. Second-order polynomial fits are used to obtain trend lines. Unlike the response that has been given in year three, students have reported their learning has improved when they spent more time on quizzes and coursework. The data suggests that students learning would not enhance by spending excessive time on quizzes which can be used as a guideline on using quizzes effectively in an engineering subject. The trend for exams is not conclusive and although it is decreasing at the beginning, it increased for those who

spent a longer period of time revising exams. This may be down to the high-stress environment surrounding exams and the fact that they occur during one period at the end of the year, whereas coursework and quizzes are more continual forms of assessment which students can do over a longer period. However, an exam was only used in one of the modules as a means for assessing students and data availability was limited.

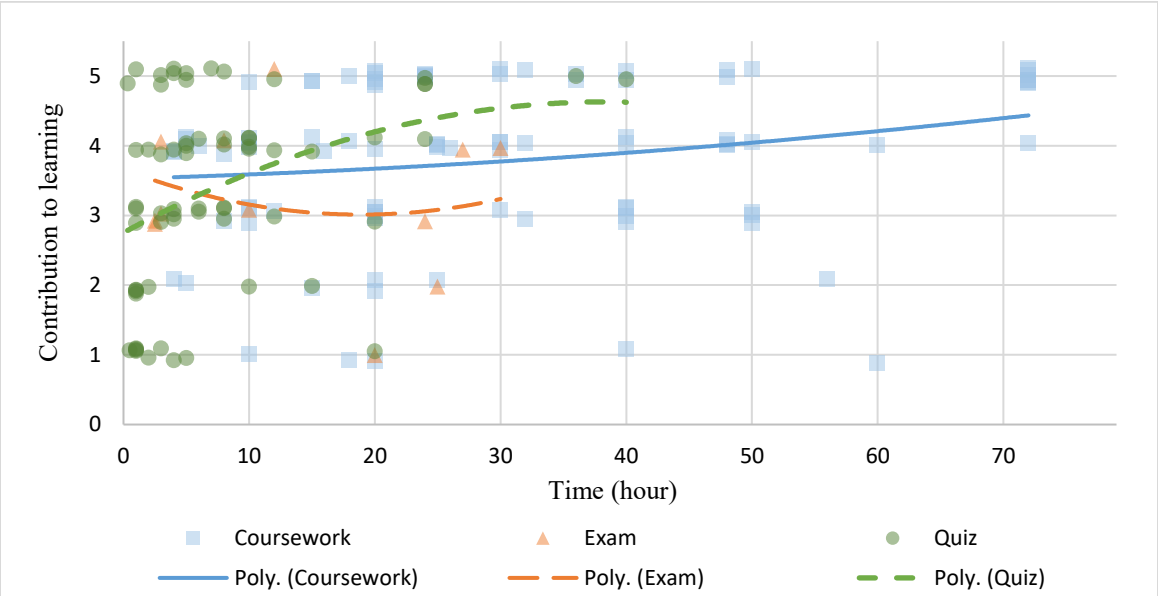

Figure 4: Contribution of different forms of assessment to students' learning as a function of the time spent on them for year two students.

4. CONCLUSIONS

The third and second year students of UCL Mechanical Engineering were surveyed in the 2020-21 academic year to collect information on how much time they spent on completing different forms of assessment and their perception of how completing those assessments has helped them to learn the subjects. The research was conducted as a response to changes in the teaching and assessment method due to the Covid-19 pandemic where the teaching was online and students were away. Year three students reported the preparation and taking exams as the most time-consuming type of assessment which helped them to learn better when they spent more time reviewing for exams. They spent less time on completing each coursework than each exam but overall, completing coursework took most of their time since most of the modules that are considered are assessed by courseworks. Online quizzes were used to engage students in the learning process and their workload was smaller than other forms of assessments. However, students that reported spending more time completing quizzes and coursework felt assessments did not help their learning.

In year two, students spent most of their time completing coursework which was also more than the duration that third years used for coursework. When year two students spent more time on coursework and quizzes, they reported the assessment contributed to their learning better. The difference between students' perceptions of how completing an assessment contributes to their learning in the third and second years can be due to the differences in the difficulty level of

modules. Third-year modules are advanced topics and when students struggle in learning them, spending more time completing a summative assessment did not help them to learn better. This emphasises the requirement of including formative assessments more widely where feedback can help students to learn without the pressure of a summative assessment. This also can be due to the third-year students' experiences. The main method of assessment in the years before the pandemic was final exams (which were replaced by an alternative assessment in 2019-20). Third-year students did not have the experience of multiple summative assessments in the forms of quizzes and coursework, while year two students only have been assessed via coursework at the end of their first year of studies due to the cancellation of their first-year exams because of the Covid-19 pandemic.

5. REFERENCES

Ababio, B. T. and Dumba, H., 2013. The Value of Continuous Assessment Strategies in Students' Learning of Geography in Senior High Schools in Ghana, *Research on Humanities and Social Sciences*, 3 (17), 71-77.

Hernández, R., 2012. Does continuous assessment in higher education support student learning? *Higher Education*, 64, 489-502.

Iqbal, M., Samiullah and Anjum A., 2017. Effect of Continuous Assessment Techniques on Students' Performance at Elementary level. *Bulletin of Education and Research*, 39 (1), 91-100.

López-Pastor V. M., 2011. Formative assessment strategies and their effect on student performance and on student and tutor workload: the results of research projects undertaken in preparation for greater convergence of universities in Spain within the European Higher Education Area (EHEA), *Journal of Further and Higher Education*, 37 (2), 163-180.

McTighe, J. and O'Connor, K., 2005. Seven Practices for Effective Learning - Educational Leadership. *Educational Leadership*, 63, 10-17

Olubukola, A. and Bankole, O., 2015. Reducing Examination Malpractices in Nigerian Schools through Effective Continuous Assessment (CA) Techniques as an Alternative to One-Shot Examination in Osun State, Nigeria. *American Journal of Educational Research*, 3 (6), 783-789.

Onuka A.O.U. and Onabamiro A.T., 2010. The effect of formative test, individual assignment and group assignment on student achievement in Junior Secondary School Mathematics, *International Journal of Educational Leadership,* 3 (3), 274-281.

Shepard, L.A., 2000. The Role of Classroom Assessment in Teaching and Learning. *CSE Technical Report 517*, University of Colorado at Boulder. Url: https://cresst.org/wp-content/uploads/TECH517.pdf

William, D., 2004, Assessment and the regulation of learning, *NCME Conference,* Sandiego, CA April 2004.