

## Using pass/fail final exams in response to the Covid-19 pandemic

(Example case study for Numbas user meeting, Spring 2022)

Chris Graham and Christian Lawson-Perfect, Newcastle University

As teaching moved online due to the Covid-19 pandemic in the Spring of 2020, Universities were forced to make difficult and swift decisions on the format of final exams due to take place that summer. Decisions balanced the importance of valid assessments with unprecedented considerations for the wellbeing and individual circumstances of students. At Newcastle University, the decision was taken to replace most exams with an online, open-book equivalent, available to students for 24 hours. Although not an ideal format, particularly for mathematical subjects, the decision allowed for a relatively normal process for moderation and the award of degrees.

In the School of Mathematics, Statistics & Physics (MSP) at Newcastle, a more relaxed format was considered for stage one assessments, based on the rationale that the exam marks do not contribute to students' final degree classification. A similar argument was used elsewhere to motivate different treatment of stage 1 students. For example, University College London replaced exams with a short piece of reflective written work, whilst others, including the University of York, cancelled stage one exams altogether. Discussions in MSP about how to proceed with stage one exams resolved that there was a single essential requirement: that students were adequately prepared for stage two of their degree programme. The decision was made to use Numbas as the platform for 'Pass/Fail' assessments, to ensure that students met a minimum standard for progressing. Assessments were delivered through our bespoke lock-down app, launched through our Canvas VLE.

Exams for 13 modules across stage one of our mathematics and physics degrees were prepared, each with 100% automatically marked Numbas content. A team of postgraduate students was employed to help develop exam content, paired up with module leaders and supported by MSP's E-Learning Unit. Whilst many of the modules already used Numbas extensively for formative and summative assessment, others required specific code development. Modules involving proofs relied heavily on banks of multiple choice questions, where students were presented with a subset of many possible questions, including fill-the-blank proofs (figure 1). We tried and rejected several models for assessing proofs.

$$f(c) - f(b) = \lambda(c - b) \implies \frac{f(c) - f(b)}{c - b} = \frac{f(b) - f(a)}{b - a}.$$

D

This shows that  $f'(c_1) = f'(c_2)$ .

E

Indicate where each of the removed lines should go.

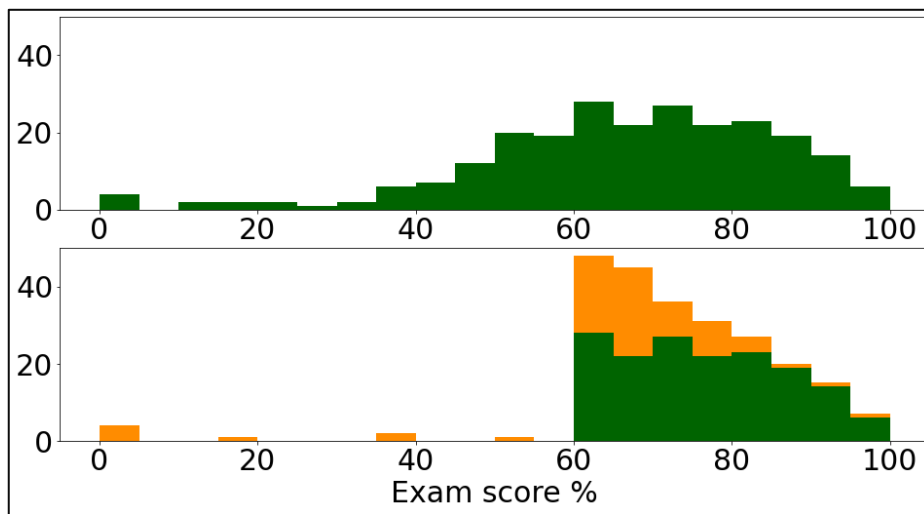
A   B   C   D   E

The equation of the line  $L$  is  $y - f(a) = \lambda(x - a)$  for  $\lambda = \frac{f(b) - f(a)}{b - a}$ .

*Figure 1: A proof question with missing lines. Students are asked where the removed lines should go. The question is randomly chosen from a set of similar questions and its content itself (i.e. the missing lines) is randomised.*

Students typically had between 3 and 6 exams, depending on degree programme. Students were able to access their exams any time in the three-week exam period and could attempt as many times as they liked to achieve a pass mark of 60% in each exam. This was significantly higher than the normal pass mark of 40% for our modules; it was chosen to match the typical median mark for our exams.

At the end of each attempt, students were shown a feedback message indicating their pass/fail status and were given information on where marks were allocated, whilst correct answers and solutions were hidden throughout.



*Figure 2: Exam scores for students in one of the pass/fail modules. The top pane shows the distribution based on each student's first score. The bottom pane shows students' final scores, with the lighter/orange scores representing students who did not pass on their first attempt.*

The results illustrated in figure 2 for a single module, multivariate calculus, were typical of most of the exams. The top pane, illustrating the first attempt score by each student, has a remarkably similar distribution to a normal exam. The second pane represents the final score, based on each student's best attempt, with most continuing until they reached the 60% pass mark. One advantage to this format is that students in the 40-60% range, who would have passed if this was a single sitting exam, were forced to improve their score and hopefully in turn will have consolidated their knowledge of the module content to better prepare for stage two.

Student feedback was mixed. Many reflected kindly on the flexible format,

*"the ability to redo the test and know how I've done immediately after was much less stressful than the 24 hour test"*

Whilst most agreed that the exams prepared them well for stage 2, others felt the format was not rigorous enough

*“Didn’t really force me to revise for the exams as I could look at notes and examples and past papers which didn’t really help me understand the content properly”*

Most negative comments related to the lack of method marks, in particular for one exam which had questions where a single answer box was worth as much as 16% of the total mark. This is certainly something to avoid in future high-stakes assessments,

Reflecting on the format, the pass/fail assessments served a particular purpose very well: to lift students to a particular standard. The non-continuation rate for students fell, as might be expected, to 10% from a long term average of 12%, suggesting that some students who may not have passed in a normal academic year were permitted into stage two. Given the unprecedented impact on personal lives in the first months of the pandemic this seems preferable to the opposite result. In addition, there has been no noticeable impact on continuation rates for the same cohort from stage two to three in the 2021 academic year.