

Using online assessment to provide instant feedback

Prof Niels Walet and Dr Marion Birch

School of Physics and Astronomy
University of Manchester
Manchester M13 9PL
(Niels.Walet,Marion.Birch)@manchester.ac.uk
<http://theory.physics.manchester.ac.uk/~mccsnrw>

Abstract

For the past five years we have been using a method inspired by the just-in-time-teaching approach to teach our first year Newtonian mechanics class, which typically contains 240-280 students. This replaced a standard course, that did not engage the students sufficiently. With such a large cohort, it is difficult to mark students' assessments quickly, and give the required feedback in a timely manner. We discuss how we capture students' marks and common difficulties in a short period, so we can plan an example class addressing these.

We explain why we have chosen a commercial e-assessment tool. We discuss the types of assignments available, especially those which allow for algebraic answers, and show which types of assessments were found most effective, and describe the underlying reasons. We will specifically discuss the steps that were needed to take to create a positive experience for the students—our initial experience was quite negative, and influenced the students' engagement.

We discuss in detail the important lessons we have learned about the mechanisms of quality control that underlie the use of online mathematical assessments. The core step in our method is to use student answers to test the quality of questions. The way of performing quality tests is highly non-trivial and time-consuming, but is crucially important. Finally, we shall discuss further developments that would make this type of rich assessment more useful to the students.

1. Introduction

Our Dynamics course, a course in which we teach Newtonian mechanics in our first year, was seen by students as slightly tedious and rather difficult. Since it is in this course that the students first meet the combination of mathematics and physics that is largely absent at A level, an effective course should instill the correct values for their university studies. In the year 2007 we redeveloped the course to use a method inspired by the just-in-time-teaching approach [1]. In this course we also employ interactive teaching techniques based on clickers [2], which we have reported on elsewhere [3], and combine it with substantial e-learning materials, including talklets [4] discussing important concepts, and physlets [5] for simulations. At the same time, we also wanted to give the students an opportunity to take charge of their own learning, by providing them with ample practice assessments with rich feedback. There is a difficulty combining all of these aims, especially since we have class size ranging from 240 to 280 students. One of the key elements of the combination is the use of electronic assessment

Even though we saw a rise in our students' achievement in exams [3] already in the first year of implementation, their initial feedback was not totally favourable. Students particularly commented on the use of e-assessment as a drawback to the course. At this point we conducted a detailed review of the e-learning material, which we shall discuss in the remainder of this paper. At the same time we introduced clicker technology into the course.

2. Just-in-Time teaching

The Just-in-Time Teaching (JiT) was developed in the USA by Novak et al [1], and is an interesting example of a student-led approach to learning. The idea of the approach is that the content of an example and/or follow up class is decided at the last minute and that this class is tailored to specifically address among the concepts the students have just studied those that they have misunderstood or are having difficulty with.

In the standard JiT approach the students are required to do some preliminary self-study. Having completed the self-study, they undertake an on-line assignment which assesses their understanding of the new material. The students submit their on-line assignments just a few hours before the class enabling the instructor to obtain an appreciation of the students' difficulties and misconceptions. He/she is then able to structure the content of his/her class according to the needs of the students. The class preparation has to be done at the last minute before the class is due to be delivered, hence the title 'Just-in-Time Teaching'.

The Just-in-Time approach was developed with and for relatively small class sizes, whereas we have around two hundred and fifty students. There also is a difficult cultural issue: the notion of pre-lecture study is quite alien to most UK students, so initially we adopted a slightly modified JiTT approach.

The Dynamics lecture course runs for 11 weeks. We start with one 'overview' lecture at the beginning of each week, followed by a Just-in-Time Response and Problem (JiRP) session at the end of the week. The overview lecture is used to 'set the scene' for the week, but does contain some conceptual questions based on the material being discussed and potentially some pre-reading. The students are expected to follow up the overview lecture with several hours of self-study using a set of e-learning material provided on the university's Virtual Learning Environment. Once they feel reasonably confident that they had understood the material, they are required to do an on-line assessment. This has to be submitted by 2.00am each Friday morning. The results of the assessments are reviewed by staff early Friday morning and preparation undertaken for the JiRP session which was held in the afternoon. The students are split into four groups for these sessions, each run by one member of staff and a postgraduate helper.

3. The role of assessment

The discussion above shows how important the role of the assessment in the middle of the week is to the JiTT approach. We thus need to find a tool that is flexible enough to allow assessment of rich mathematical content, preferably geared towards Newtonian Mechanics. There are very few systems of sufficient richness; in a general mathematical context we are aware of WebWork [6], Stack [7], LON-CAPA [8] and MapleTA [9]. These were not considered at the time of deployment of our system, but if we were to start from scratch, we might consider them today. We instead opted for a mature and feature-rich system that already came as a companion to our core first year textbook [10], mastering physics [11]. This is an offshoot of the MIT CyberTutor project [12], and as such has one of the longer pedigrees in this arena.

Mastering physics contains a number of problem types; from simple “order the pictures” and multiple-choice questions, to questions involving simulations, numerical answers with random variables, and complete algebraic answers. It is the last two we make most use of, even though we also use a few of the other types in our assessment where appropriate. There are two styles of these problems. Initially we used a mix of both tutorial and homework problems to assess our students; tutorial problems give hints and feedback, homework problems (i.e., end of chapter problems) are just marked without feedback. As one can see by typing “MIT mastering physics” into a search engine, there are real difficulties with motivating students for the homework style problems. The feedback we got in the first year of usage was rather negative and mirrors what one can find on the web, so the next year we opted to use the flexibility within the system to recast all questions in the tutorial mold. At the same time we removed hints that we considered too obvious, so that students cannot attack the problems with no background reading at all. The result of this is a very positive reaction of the students to Mastering physics—so much so that there have been requests to adopt it in other courses. Hand-in rates for homework exceed those for tutorials. There is some indication of game-playing in the way students answer questions, but on balance the results seem good.

A key difficulty with giving answers to questions with mathematical content is entering mathematics on a computer in a consistent way: Mathematical notation is surprisingly ambiguous, so whatever system we use, we need to set some ground rules as what textual form represents a product, or a power, etc. We can use graphical interfaces to enter such equations; Mastering physics has its own tool, and the open source systems tend to use a java-based plugin called dragmath. Of course students need to be trained on first usage, so that they can use such systems effectively; nevertheless, we always get a few complaints from students who claim they have been marked incorrectly, whereas they have in reality failed to follow standards.

4. Feedback and malrules

Tutorial questions add two important layers to standard assessments: the option to open and use hints (which can contain further questions where we can ask the students to construct partial answers), and a rich layer of feedback based on recognised flaws in their answers. Authoring hints is relatively straightforward, even though quality control requires some additional work (see next section). Feedback is more difficult.

The specific nature of giving rich feedback in questions where the answers are of an algebraic and algorithmic nature, clearly lies somewhere between giving feedback on multiple choice questions and on lengthy full text answers. In order to undertake this task, we must develop a model for the answer we expect, and analyse the received answer against that. Usually, systems with this ability built in will by default look for unknown variables and other simple syntax errors, as well as rounding errors, but suppose a student has differentiated rather than integrated her result—how do we provide sensible feedback?

This brings us to the important area of *malrules*. A malrule is misconception that can be captured as applying an incorrect algorithm. This can be a simple mistake, but sometimes it can be the sign of a deeper misunderstanding; these can also be linked to a threshold concept that has not been mastered yet. They usually apply to problems which require a slightly higher level of understanding and analysis.

Once one considers solving a physics problem in this algorithmic manner, we can try to identify the algorithm from the result inputted by the student. With simple enough problems, this can be quite successful, and thus we may give our students feedback such as “*You have differentiated ,, with respect to time, but since you are solving a differential equation, shouldn't you integrate instead?*” Such feedback, with a chance to resubmit an answer for partial credit, can be very beneficial to students. It seems to contribute to their understanding.

5. Information capture

Most VLEs provide some level of aggregate summary of the answers given by students to tests, often with the ability to drill down to the level of individual answers by individual students. More complex question types require that the system provides detailed useful reports about the usage of the system, as well as various levels of drill-down into groups of answers and individual answers in order to ensure validity of the assessments.

Masteringphysics has some of the best analysis tools we have met, allowing to identify all the different answers given by students with frequencies and feedback provided for each such answer. One of the other features that has proven very useful is the ability for students to rate the difficulty of each problem, as well as to provide comments on an assessment. Together with time spent and marks obtained, these provide a good basis for a validation of the assessments.

6. E-learning quality control

One of the important lessons we have learned is about the mechanisms of quality control that underlie the use of this type of algebraic assessment. There is a body of knowledge about developing and checking the validity of MCQs, but little has been published about validation of more complicated questions.

For our case, we need to analyse both the quality of hints, and the feedback students get on their incorrect answers. This requires very specific reporting ability in the assessment software—something commercial VLEs are still working towards. Fortunately, mastering physics does have a very good set of tools for this purpose.

Let us first discuss the easier of the two: hints. The system allows us to see what fraction of the students opens a certain hint, and also how many students answer associated questions correctly. We find that the fraction of students using hints is a clear indication of the need for a hint. Together with their difficulty ratings and time spent, this a very useful indicator of whether hints are relevant. Usefulness is often obtained from the written feedback they can give at the end of the assessment. All together this allows for a considerable fine tuning of the choice and provision of hints, even we are not completely satisfied: there is room for further analysis.

The second problem, providing quality control for feedback looks daunting at first sight. The core step in our method is to use student answers to test the quality of questions, which in hindsight seems obvious. The way of performing quality tests is highly non-trivial and time-consuming. What we do is the following; the system allows us to see an amalgam of a question and all the answers plus frequency and feedback given to students. We then need to consider whether this was appropriate. In most cases, we see that the feedback given and the expected malrule applied match, and there are few if any answers with no feedback. What we then do is study all questions, and see where we have insufficient feedback, and how that could be matched to malrules or mistakes. When we first did this analysis, we were able to provide such malrules wherever they were required. We also noted some simple mistakes (e.g., a lower case variable in the malrule, where an uppercase was expected), some of which originated in the underlying question banks.

Such an exercise adds substantially to the costs of deploying such assessments. The first two times we re-authored questions and malrules based on student input it took about 40 hours for 50 questions. After five years, this has come down less than a day.

7. Outlook

The detailed type of analysis required to ensure quality of e-assessment in a mathematical context means that as they get embedded in VLEs as new question types, we should not lose track of the reporting tools required to ensure that assessments are valid and of sufficient quality. Modern VLEs have some, but they are often rather limited.

A special purpose commercial product such as mastering physics sets a high standard, and we can only hope that other products will compete at this high level. This is the only way that we will be able to use similar high-quality material in many other settings.

8. References

1. Novak, G. M., Paterson, E., Gavrin, A. D. and Christian, W. (1999) *Just-in-Time Teaching: Blending active learning with web technology*, New Jersey: Prentice Hall.
2. Mazur, E. (1997) *Peer Instruction: A User's Manual*, New Jersey: Prentice Hall.
3. Birch, M. and Walet, N.R. *An Integrated Approach to Encourage Student-Centred Learning: a First Course in Dynamics*, New Directions in the Teaching of Physical Sciences, Issue 4 December 2008 pp. 21-26, HEA Physical Sciences Centre, ISSN 1740-9888
4. Levitt, M. (2004) *Talklets: A New Way to Teach*
5. (<http://csb.stanford.edu/levitt/Talklets.html>) (accessed February 2011).
6. Christian, W. (2005) *Physlets* <http://webphysics.davidson.edu/Applets/Applets.html> (accessed February 2011).
7. The Mathematical Association of America (2011) *WeBWork: Online homework for math and science* <http://webwork.maa.org/moodle/> ((accessed February 2011).
8. Sangwin, C (2010), *Stack* <http://www.stack.bham.ac.uk> (accessed Dec 2011)
9. LON-CAPA (2011) The LearningOnline Network with CAPA <http://www.lon-capa.org/> (accessed Feb 2012)
10. MAppleSoft (2012), Maple T.A., Testing, Grading and Evaluation Software, <http://www.maplesoft.com/products/mapleta/>
11. Young, H. D., and Freedman, R. A. (2010) *Sears and Zemansky's University Physics with Modern Physics*, 13th Edition, San Francisco, USA: Pearson Education, Addison Wesley.

12. Mastering Physics (2012) (<http://www.masteringphysics.com>), Pearson Education publishing as Addison Wesley (accessed Feb 2012).

13. Pritchard, D. E., & Morote, E.-S. (2002). *Reliable assessment with CyberTutor, a web-based homework tutor*. In World Conference on E-learning in Corporate, Government, HealthCare, and Higher Education, Montréal, Canada. E-Learn.