Quantum: tests worth teaching to

Status report October 2017

1   Background

[Project Quantum](#) aims to help computing teachers check their students' understanding, and support their progress, by providing free access to an online assessment system. Distinctive features of Quantum (do read the [white paper here](#)) are:

- **Formative**. Quantum is focused on frequent, low-stakes, formative, diagnostic assessment to support learning (in contrast to high-stakes summative assessment).
- **School-led, crowd-sourced.** Teachers both use the corpus of questions on the system and upload questions of their own.
- **Open**. Quantum uses a free, online platform, [Diagnostic Questions](#). Moreover, the questions themselves can be re-used by other platforms; and anonymised data will be available to researchers.
- **Evidence-driven, research-led.** Quantum partners include two of the leading assessment experts in the nation, Tim Oates (Cambridge Assessment) and Robert Coe (Durham Centre for Evaluation and Monitoring).   The CEM contribution will be to provide quality control for the crowd-sourced questions, by analysing the data from thousands of students doing thousands of questions. No one has ever done this before.
- Research and reality**.** The project combines two goals
    - **Reality**: being immediately useful to practising computing teachers. They have a crying need for high-quality assessment material, and Quantum will produce this, quickly. We aim to cover both primary and secondary.
    - **Research**: no one has tried to crowd-source assessment items, and then use data to evaluated and improve their quality. If we can make this work, the results will be useful for all subjects in any country. We aim to change the world!

The initial project is generously funded by Google, Microsoft, and ARM, over two years, starting April 2016.   The main project partners are:

- Computing at School (CAS)
- Durham Centre for Evaluation and Monitoring (CEM)
- Cambridge Assessment
- Eedi / Diagnostic Questions

This status report summarises our progress over period April-September 2017, and current status.

2 Headlines

Quantum is proceeding apace, on three fronts:

- **Content**. The computing content is developing rapidly (Section 3). We now have over 3,000 questions in the system, and that is enough to be genuinely useful. Alongside developing more content, we will now focus on increasing usage.

- **Analysis and quality control**. The unique feature of Quantum is our ability to analyse data from thousands of students answering thousands of questions, to provide quantitative, data-driven feedback to authors about the quality and effectiveness of their questions (Section 4).

- **Platform**. We are designing and implementing changes to the Diagnostic Questions platform itself, in direct response to the needs of the first two strands; Section 5 elaborates.

3 Developing computing content (CAS)

3.1 Questions

The number of Computing questions in Quantum is rising steadily. We have just exceeded 5,000 questions, and have at least another 1,000 in the pipeline. This is an increase of over 2,000 since April.

It is already a usable and useful resource for computing teachers wishing to avoid reinventing the wheel for low-stakes, formative assessment of their pupils' knowledge and understanding in computing.

This chart shows our progress over time.

Question total by month
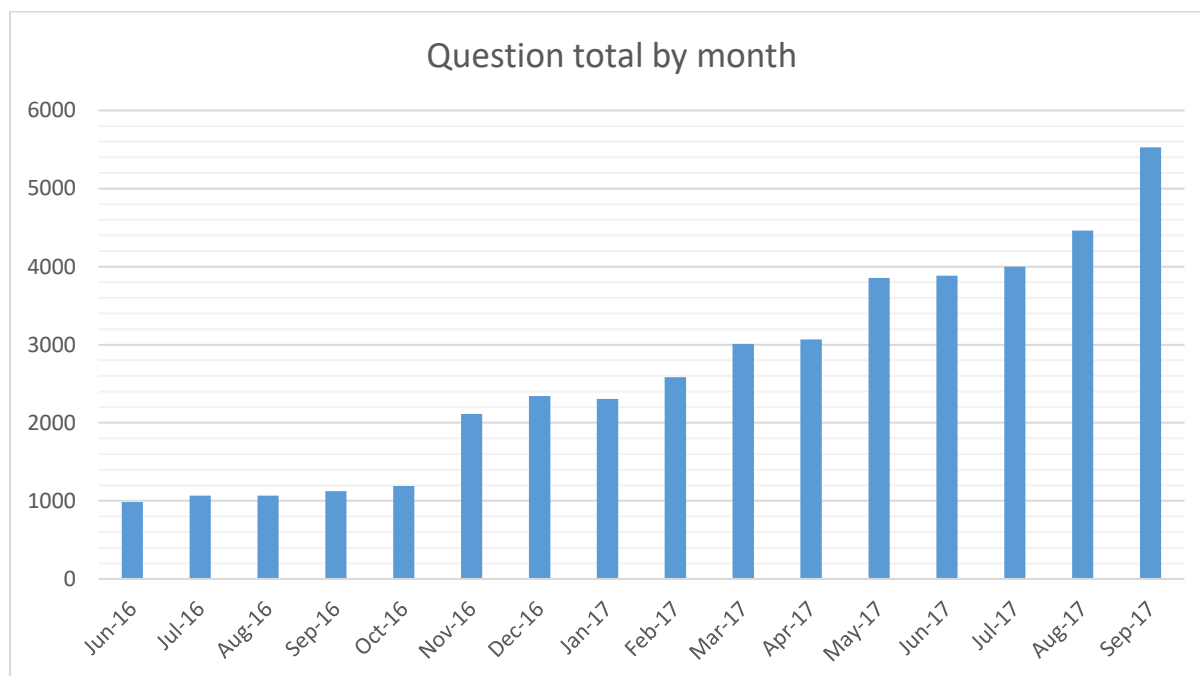
3.2 Quizzes

A huge corpus of questions, even if they are well-tagged, is somewhat daunting for a teacher. In our April 2017 report we wrote

A priority for the next phase of the project will be to gather together questions into key stage / topic based quizzes, to make it easier for teachers to make use of the materials developed to date.

In this six-month period we therefore focused on gathering **questions** into **quizzes**. Quizzes are easier for teachers to use, because each quiz is a manually-chosen set of questions, aimed at a similar ability range, on a particular topic. Moreover, we have gathered quizzes into sets of three on each topic:

- a baseline quiz to use before starting the topic
- a checking quiz to use right and the end of the topic
- a revision quiz to use later, to check retention

We now have over 250 quizzes, gathered together on the DQ Quantum home page: https://diagnosticquestions.com/Quantum.

We have also developed several longer "baseline" quizzes, that teachers can use to characterise their students level of prior knowledge at the start of various school years. We now have

- Five baseline quizzes from TLM
- Four baselines curated by CAS
    - End of KS2 Baseline https://diagnosticquestions.com/Quizzes/Go#/57435
    - Year 7 Baseline https://diagnosticquestions.com/Quizzes/Go#/59118
    - Year 8 Baseline https://diagnosticquestions.com/Quizzes/Go#/59119
    - Year 9 Baseline https://diagnosticquestions.com/Quizzes/Go#/59120

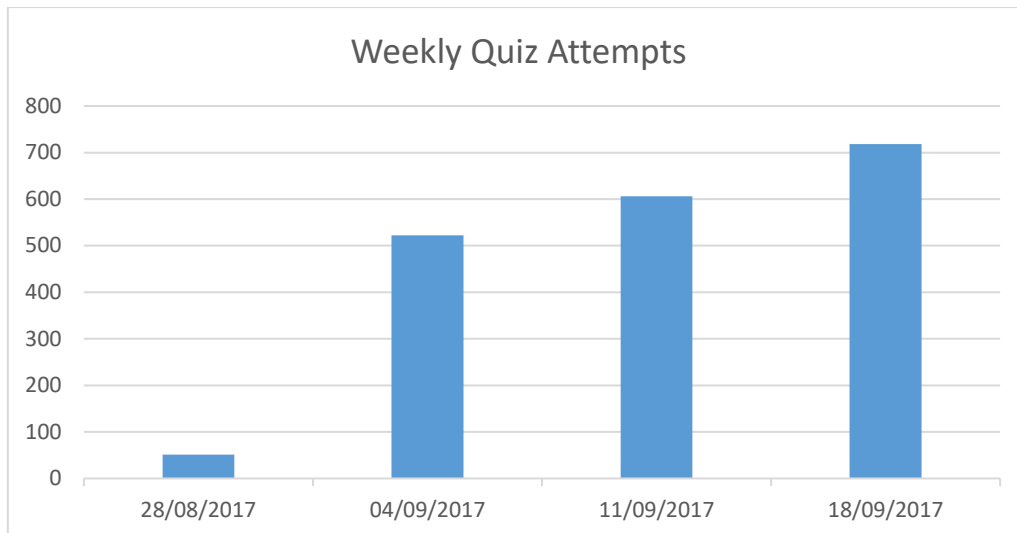Each baseline quiz has 50 questions, or else 3 shorter sets of 16 or 17 questions.

This represents a very significant new investment that we hope will help drive adoption.
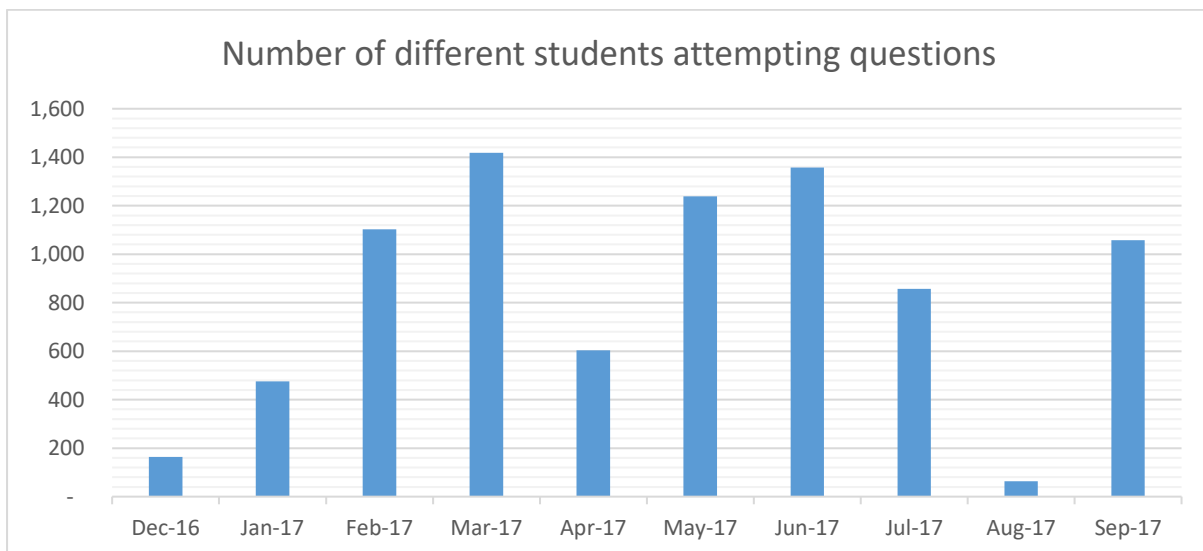
3.3   Usage

The number of questions actually answered by students is also rising rapidly, and has nearly doubled since March:
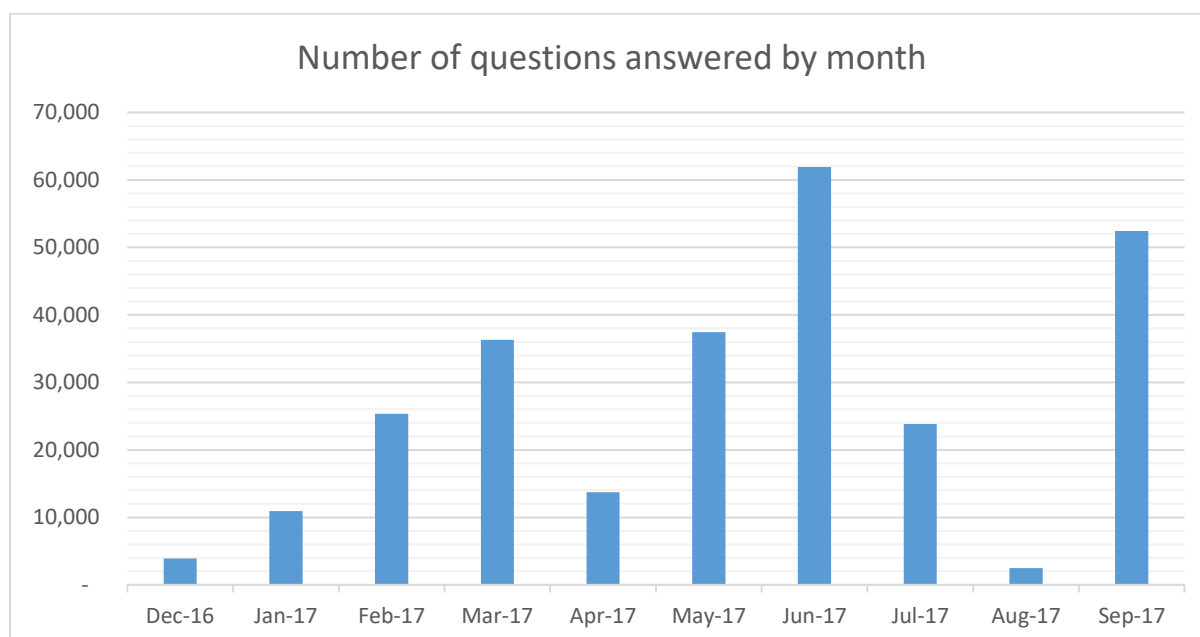
The large dips are attributable to the school holidays, and September is a partial month. Here is weekly data on the number of quizzes attempted in the first few weeks of the autumn term.

Weekly Quiz Attempts

A bar chart titled "Weekly Quiz Attempts" showing the following approximate values:
- 28/08/2017: ~50
- 04/09/2017: ~520
- 11/09/2017: ~605
- 18/09/2017: ~720

On the other hand, the numbers of students involved does not appear to be rising:

Number of different students attempting questions

A bar chart titled "Number of different students attempting questions" showing the following approximate values:
- Dec-16: ~170
- Jan-17: ~480
- Feb-17: ~1,100
- Mar-17: ~1,420
- Apr-17: ~600
- May-17: ~1,240
- Jun-17: ~1,360
- Jul-17: ~860
- Aug-17: ~70
- Sep-17: ~1,060

3.4   Coverage

Although the Quantum project aims to cover the whole of school-level computing, based on the English national curriculum plus GCSE and A Level, the current collection of questions is focussed more heavily on computer science (CS) than either information technology (IT) or digital literacy (DL), although the coverage of these latter two elements is certainly improving. Of the 5,529 questions,

- 3,714 are tagged as CS/foundations, including
    - 2,007 relating to systems (which includes the popular categories of hardware, binary representation and binary arithmetic
    - 1,332 are on programming
    - 918 on computational thinking
- 963 are on IT/applications
- 764 on DL/implications.

The questions cover GCSE computer science and Key Stage 3 more extensively than A Level or primary.

3.5   Publicity and uptake

Although you can never have enough questions, we believe that we now have a large enough corpus, well enough organised, that it should be directly useful to school teachers. So our priority for the next six months is to **drive adoption**.

It is hard to know how best to do this. We have presented Quantum at a number of regional conferences, and at the national CAS Conference in June, and adoption is rising. But it is still well below what we would like.

One strategy we plan to try is to develop one or more Schemes of Work for Computing. In this context a Scheme of Work is a linked series of quizzes, that teachers can use for a whole term or year, to guide their teaching and assessment. Eedi have found this to be very effective for Maths. We plan to develop one for computing during the next six-month period.

4   Data analysis (Durham CEM)

The unique feature of Quantum is the use of quantitative methods to offer evidence-based feedback about the quality and effectiveness of assessment items. Our hypothesis is that this feedback will drive adoption; and adoption will produce more data that will improve the reliability of the feedback, in a virtuous cycle.

4.1   Audience and priorities

There are two audiences for the quality feedback

1. **The author of the question**. The goal is to produce actionable ideas that will help the author improve the question, or to write better questions in future.

2. **A teacher considering whether or not to use the question**. The goal is to give the teacher guidance about
    - Whether the question is **effective**; that is, whether it discriminates between high and low ability students.
    - What **ability range** the question targets. An excellent question for primary children may be useless for A-level students.

Up to now CEM has concentrated on (1). Quite a bit of progress has been made (see the April 2017 report for example), but it is hard to point to solid achievements:

- Turning Rasch analysis data into actionable feedback is hard for an expert to do, let alone an automated system.
- The question authors are not experts. If they were, the task would easier. But they aren't.
- Authors are relatively few, and teachers are extremely busy. That makes it difficult to find enough authors who have time to engage with CEM in improving the quality of the feedback mechanisms. (We found just two such authors in this period.)

So we are now shifting out priority to focus on (2):

- For every author (producer) there are hundreds of teachers who just want to use the system (consumers), so the population is much larger.
- The actionable guidance is much easier to characterise. The main questions are:
    - Is this question any good?
    - How hard is it, relative to the ability of my students?

4.2   Automation

A specific goal of Quantum is to automate the feedback process. Up to now, the data analysis has all be manually driven. But by March 2018, and hopefully sooner, we will automate the process. Specifically:

- All questions in Eedi's platform, for which there is enough data, will get automatically-updated feedback about question quality and ability range, that will be visible to all users.

4.3   Communication

Mark Frazer spoke at this year's Computing at School North East Regional Conference in Newcastle, on 6th July 2017 where he presented a brief introduction to Project Quantum and the work that has been conducted so far.

Approximately thirty delegates attended the presentation, with only three of them having a prior awareness of the project, or the Diagnostic Questions platform. Speaking at the conference was therefore a good opportunity to tell delegates about our work and, more importantly, encourage them to sign up to use DQ and, ideally, contribute some computing items to the bank.

The current DQ format and feedback was explained and a short demonstration of the platform's functionality was given, which was well received. Teachers were particularly interested in the potential for diagnosing misconceptions and the facility for students to qualify their responses with a short comment triggered some interesting discussion about the potential value of this information.

Following this, a brief overview of the investigative work was given with emphasis being placed on the dashboard created by John and Sarah before the distractor analysis was introduced. Delegates liked the idea of the simple dials and descriptive statistics for quality, difficulty, percentage correct and popularity. There was also interest in the power of the distractor analysis plots, however, it should be noted that there was one Rasch enthusiast in the audience (not planted) who was influential in setting the mood of the audience! A discussion with the audience revealed that they were interested in the information visualised in the distractor plots, but there is a need to provide clearer guidance on their interpretation.

The presentation was concluded by providing potential item authors with some basic guidance on writing multiple choice questions.

5   Developing the platform (Eedi)

5.1   Integration with School Management Information Systems

Diagnostic Questions is heavily used by mathematics teachers across the country. We identified a bottleneck in usage when teachers were required to add their students, so many teachers were simply using the questions at the front of the classroom and not assigning quizzes for their students to complete.

To address this bottleneck we have added support for syncing teachers, students, classes, departments and year groups directly with school Management Information Systems. The benefit of this integration for the Quantum project is that when schools choose to integrate in this way all the computing teachers and classes are also imported which reduces the barrier to entry for computing teachers.

5.2   Eedi

The big news from Diagnostic Questions is that we are (slowly) rebranding as Eedi. The new site can we found at https://eedi.co.uk (you can log in with your Diagnostic Questions credentials). Initially Eedi is focused around schemes of work, it is makes it easier to adopt, edit and assign schemes to

classes. Support will soon be added for students and parents. We only have schemes of work for mathematics at the moment but we are hoping to develop a scheme of work for computing.

5.3   API

As part of the development of Eedi we have been building an API for Diagnostic Questions (Eedi is being built on a completely new stack). So far the API supports membership and scheme of work functionality, and we are now adding support for answering quizzes.

Partly to support Quantum, a subset of the API is being made publicly available to support integrations with MOOCs. This will make it possible for partners in Quantum to use the corpus of questions without necessarily using Eedi's delivery platform. Crucially, however, we will still be able to gather the data from student answers.

5.4   Quantum Homepage and Curated Collections

The www.projectquantum.org page which was created previously was a static site which made it difficult to add content. We have now removed the static site and pointed www.projectquantum.org to http://community.computingatschool.org.uk/resources/4382 which is the home for the Quantum project on CAS.

The content team has been uploading quality questions and collating them into curated collections with the aim of making it easier for teachers to find good questions. To make these collections more visible we created a homepage on Diagnostic Questions https://diagnosticquestions.com/quantum. On this homepage you can find a summary of the project; a description of how to find computing resources on the site; and multiple quizzes covering CS Foundation, IT Applications and DL Implications.

5.5   Future

We have (finally) started the development of some metric dashboards to track usage on Eedi and Diagnostic Questions. In our roadmap we have planned to complete this by 1st November 2017.

6   Other research

Yerkhan Mindetbay began his PhD studies at the University of Southampton in September 2016, co-supervised by Dr John Woollard and Dr Christian Bokhove.   He has gradually focussed his attention towards establishing the efficacy of multiple choice questions (MCQ) in the assessment of computer science. His work will contribute to the Quantum question bank, and will also provide evidence for the future evaluation of the assessment processes. He is working closely with Durham CEM.

This section, written in his own words, summarises his progress and plans.

6.1   A focus on computational thinking

To assess computational thinking skills by means of "diagnosticquestions.com" I decided to choose several concepts such as Abstraction (identifying visualisations, identifying flowcharts for written pseudocode), Generalisation (patterns with numbers, patterns with figures, patterns in decoding), Algorithm (logic gates on circuits), and Logic (narrative approach to logic, logic with numbers and letters). Each of these mentioned subtopics will have three different 10 multiple-choice questions with the same level of difficulty, to be administered: before the course, during the course and after the course. So for each subtopic there will be one set of multiple-choice questions (30 questions).

6.2   Questions for Project Quantum

My initial plan is to prepare and upload 8x30=240 questions by the end of October, and add more lately. While preparing these questions, I will take into consideration the new Computer Science

curriculum of Kazakhstan and the new textbooks accordingly, so that my diagnostic questions would meet the requirement to measure computational thinking skills and be compatible with Kazakhstani curriculum. I am planning my piloting for late November and early December.

6.3   Question design

There are several recommendations on preparing good multiple-choice questions arising from published papers. A good quality multiple-choice question is the one that assesses not just lower-order thinking but also higher-order and deeper thinking and understanding. There are several common issues with multiple-choice questions such as poor-quality questions with implausible answer choices, factors like guessing, etc. I hope to make better quality questions; feedbacks and exam statistics such as item-total correlation and overall test scores will help us to make it possible. Item Response Theory will be used to guide the analysis the statistical data on tests and improve question quality. The validity/reliability of MCQ is an important aspect of this work.