PTAS Project Report (for REGULAR PROJECT GRANTS)

Project Title: Improving the Mathematics Diagnostic Test

Project type

B Innovation Project (introduction and evaluation of an educational innovation, usually taking a practical approach)

Principal Investigator: Dr George Kinnear

Schools/department: School of Mathematics

Team members: Dr George Kinnear, Professor Chris Sangwin, Dr Toby Bailey (School of Mathematics)

For further details, please contact: Dr George Kinnear (G.Kinnear@ed.ac.uk)

What did you do?

Diagnostic tests are widely used in university mathematics to inform staff about the level of students' basic mathematical skills, and to inform individual students of gaps in the level of assumed mathematical knowledge. The mathematics diagnostic test used by the University of Edinburgh is completed online by around 1000 students each year, as it is made available to any new student taking a course taught by the School of Mathematics.

In 2017 the decision was taken to move to delivering the test through the STACK computer-aided assessment system (Sangwin, 2013), and this presented an opportunity to rethink the content of the test. In this project, we used data on previous students' attempts at the test to evaluate its performance and suggest improvements (with much of the analysis being carried out by two undergraduate students). These changes were implemented in September 2017, and the new version of the test was administered to new students in 2017-18. Further analysis was carried out in the summer of 2018 (by a third undergraduate student), using data from this first administration of the new test to check that the changes had been successful. Our analysis also helped to inform advice to students and Personal Tutors about how to interpret the test results.

In evaluating the test, four different aspects were considered: the mathematical content of the test, exploratory factor analysis, psychometric item analysis, and the relationship of test scores with performance in subsequent mathematics courses. These approaches and our results are described in the next section.

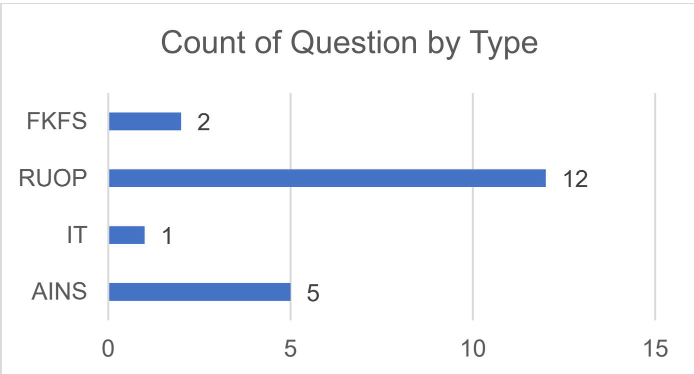
What did you find out?

Mathematical content

The mathematical content fits closely with the syllabus of the Scottish Qualification Authority's "Higher Mathematics" course (SQA, 2016) since this is the minimum qualification required for entry. The topics covered by the test are thus largely in algebra, trigonometry and basic calculus. The main analysis conducted in this project was with respect to the types of questions being asked on these topics, according to the "Mathematical Assessment Task Hierarchy" (Smith *et al*., 1996), also known as the "MATH taxonomy", which can be used to classify questions based on the skills needed to complete the task successfully. The following table summarises the different question categories:

Group A Routine procedures	FKFS: Factual Knowledge and Fact Systems COMP: Comprehension RUOP: Routine Use of Procedures
Group B Using existing mathematical knowledge in new ways	IT: Information transfer AINS: Application in New Situations
Group C Application of conceptual knowledge	JI: Justifying and Interpreting ICC: Implications, Conjectures and Comparisons EVAL: Evaluation

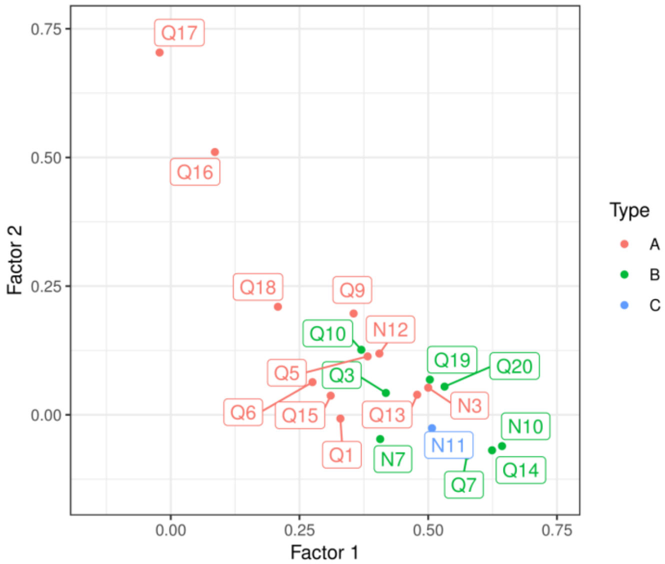
Four members of the project team independently classified the questions on the test, then met to discuss and agree the classification. The agreed classification shows that the test was predominantly Group A questions, based on "routine procedures". Additionally, there were no Group C questions.



This is perhaps reasonable given the nature of the test (to revise core skills already learned), however it did suggest that in revising the test, there should be a preference for removing Group A questions where possible, and that these should be replaced with Group B and Group C questions.

Factor analysis

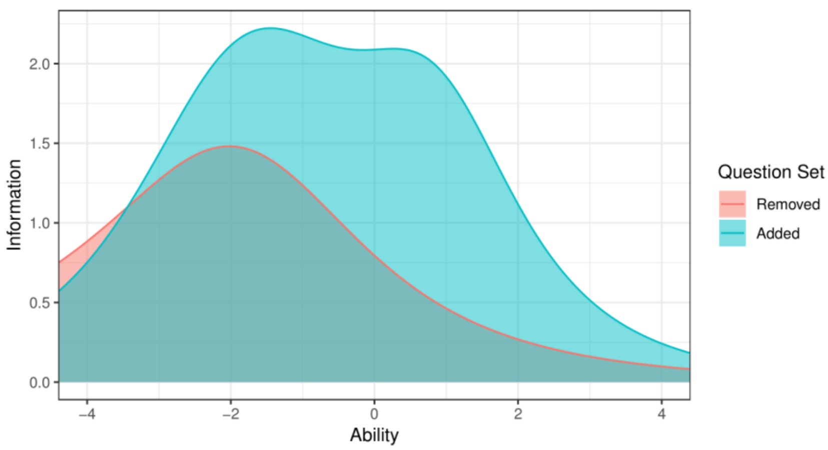
Factor analysis is a statistical technique which essentially seeks to break the test up into groups of questions where the scores in each group correlate well but the correlation between groups is lower. This is driven by student response data, but the aim is to find groupings which can be given a meaningful description, e.g. "algebra" versus "calculus". For the diagnostic test, two dominant factors were identified. Strikingly, the questions which had been classified as MATH Group B were loaded most heavily onto one factor.



Item response theory

A psychometric item analysis was carried out, providing a model of how students of various abilities would perform on each question. This gives a visual representation which shows the spread of question difficulties, and also how well each question discriminates between the most and least able students (Embretson & Reise, 2000). On the basis of this analysis, three of the least discriminating items (all MATH Group A) were selected to be replaced.

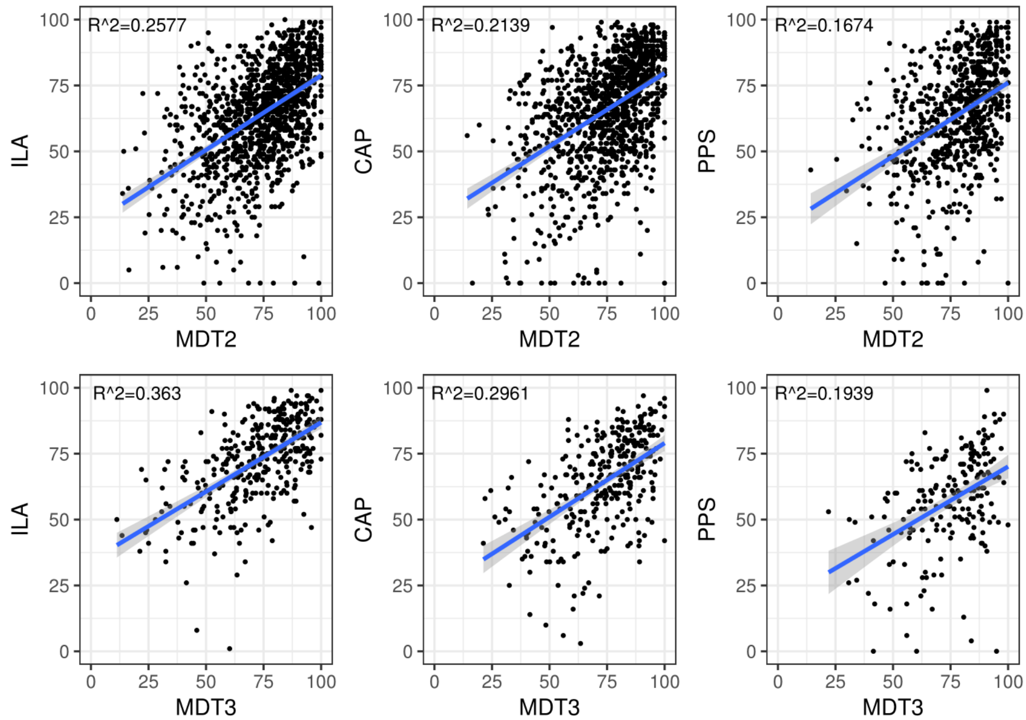
Item response theory also gives a measure of the “information” provided by each question; this quantifies how well a given question helps to distinguish between students of a given ability level, e.g. a very difficult question will help to differentiate between the most able students, but will not help much to distinguish between less able students since they will all get it wrong, so the information for this question would be mostly weighted towards the high ability end of the scale. One measure of the success of our changes to the test is given by comparing the information curves for the questions removed with those for the questions added:



The information curves show that the new version of the test provides better information about students’ abilities, particularly around the mean ability level (which is fixed at 0 on the scale).

Predictive validity

Students' test scores were compared with their performance in various subsequent mathematics courses. Since the test scores are used in practice to help inform course choice, it is important to understand this relationship. We found that the existing test (labelled MDT2) correlated moderately with performance in Year 1 Mathematics courses; for instance the correlation with the first semester course Introduction to Linear Algebra was r^2=0.2577. Moreover, we found that the revised version of the test (labelled MDT3) correlates more strongly with results in these courses; for instance the correlation with Introduction to Linear Algebra has increased to r^2=0.363.



How did you disseminate your findings?

Conference talks

"Improving a Mathematics Diagnostic Test", MatRIC Conference 2017 *Excellence in Teaching Higher Education Mathematics*, Oslo 9-10 October 2017, https://www.matric.no/events/25

"Using item response theory to evaluate a test", E-Assessment in Mathematical Sciences (EAMS) Conference 2018, Newcastle 28-30 August 2018, https://eams.ncl.ac.uk/

Conference poster

"Improving an online diagnostic test via item analysis", Mathematics Education in the Digital Age (5th topic conference of the European Society for Research in Mathematics Education), Copenhagen 5-7 September 2018, https://www.math.ku.dk/english/research/conferences/2018/meda/proceedings/posters/

Student information

Analysis from the project contributed to new guidelines for student course choice, published at https://teaching.maths.ed.ac.uk/main/undergraduate/about-your-studies/pre-honours/year-1

What have been the benefits to student learning?

The new test provides a more accurate measure of students' abilities, so will help to inform appropriate course choices – in particular, the School of Mathematics has introduced a new course in 2018/19, "MATH07003 Fundamentals of Algebra and Calculus", which is particularly recommended for students who score below 50% on the diagnostic test.

How could these benefits be extended to other parts of the university?

The use of item response theory to analyse exam or test results could help to improve the reliability of assessments in other contexts.

References

Embretson, S. E., & Reise, S. P. (2000). *Multivariate Applications Books Series. Item response theory for psychologists.* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Sangwin, C. J. (2013) *Computer Aided Assessment of Mathematics*. Oxford University Press.

Smith, G. *et al.* (1996) 'Constructing mathematical examinations to assess a range of knowledge and skills', *International Journal of Mathematical Education in Science and Technology*. Taylor & Francis Group , 27(1), pp. 65–77.

SQA (2016) 'Higher Mathematics Course Support Notes'. Available at: https://www.sqa.org.uk/files_ccc/CfE_CourseUnitSupportNotes_Higher_Mathematics_Mathematics.pdf (Accessed: 1 November 2018).

Financial statement:

This project has utilised the funding awarded to it by the PTAS adjudication committee and the Principal Investigator or School Administrator appropriate can provide financial statements showing the funding usage as and when required by the UoE Development Trusts who may require it for auditing purposes.