

Teaching Committee report to the Faculty Board

Examinations for Parts IA, IB and II of the Tripos, 2017

1 Introduction

The Committee met four times, for two hours each, to consider the 2017 undergraduate Tripos examinations. As usual we looked, for each part of the Tripos, at:

- the examiners' report;
- the external examiners' reports;
- the examiners' comments on their questions;
- the examination statistics;
- the examination papers;
- the analysis of the paper lecture questionnaires;
- the responses to the on-line questionnaires;
- the report from the CATAM assessors (Parts IB and II).

We noted with pleasure that as usual the external examiners, without exception, commented favourably both on the examination process and also on the performance of the candidates. The six external examiners' reports include comments such as:

The Mathematical Tripos in Cambridge remains one of the most ambitious Mathematics courses in the country, and the examination reflects this. (Part IA)

I am overall very happy with the examining procedure of which I was a part. The level of difficulty of the questions is appropriate, and the processes for assessment, examination and the determination of awards are sound and fairly conducted. (Part IB)

Compendium examination papers remain appropriate to the students at a world-class university. (Part IB)

As in my other two years, the standard of the examination papers remained very high and I was impressed by the performance of the best candidates. (Part II)

I gained an extremely favourable impression of the quality of the students, and of the fairness and rigour of the examination process. I am happy that Cambridge Part II mathematicians continue to represent the best of UK mathematics, and the course acts as a beacon of high standards, certainly matching or exceeding any other UK programme. (Part II)

The standards are markedly higher than at any other university where I have worked or served as an external examiner. (Part II)

There follows a summary of the points raised in the examiners' reports which the Committee believe need the attention of the Faculty Board. We have not generally highlighted points of a purely administrative nature: that is for the Chairs of this year's examiners and the Undergraduate Office to pick up rather than the Faculty Board.

2 General Matters

2.1 Errors

This year only six errors were discovered during or after the examinations, in a total of about 280 questions: no errors in Part IA; one error in Part IB; five errors in Part II. In addition, three errors were discovered before the examination and announced at the very start. Overall, these numbers are smaller than most recent years, though larger than last year.

The Part II examiners recommend that all questions are made to conform to the Faculty's typesetting conventions *in advance* of the third meeting of examiners. This should be done by the Chair (and perhaps their deputy from the other department). This would allow the examiners to focus on the mathematical content and difficulty in the meeting, rather than being distracted by issues of style.

We strongly **recommend** imposition of the typesetting conventions *before* the second or third examiners' meeting, if necessary by the Chair.

2.2 Short questions

Last year we noted that the success rate for betas on short questions was still a bit short of the target 65–70% and recommended that pressure was maintained to keep short questions both short and easy. We are pleased to see progress as shown in the following table:

	beta rate (%)			average mark			takeup in II	
	IA	IB	II	IA	IB	II	short	long
2017	62	65	59	7.1	7.2	6.7	14.6	17.1
2016	59	59	57	6.8	6.8	6.5	17.4	17.1
2015	44	52	51	6.7	7.0	6.8	12.4	15.8
2014	60	51	48	7.2	7.0	6.6	12.9	15.5

Profs. Anderson and MacKay (Part IB externals) and Prof. Johnson (Part II external) all mention the importance of short questions as a means for giving even the weakest students opportunity to demonstrate success in learning. We entirely agree, and hope that examiners will continue to push in this direction.

2.3 Comments on questions

Faculty Board has agreed that the examiners' comments on questions will now be put online in the January following. Several examiners' reports requested guidance regarding the intended length and substance of these comments.

We **recommend** that the comments should describe how the candidates performed compared with the examiners' expectation of the question, which parts were done well and what were the most common mistakes. Two or three sentences would be appropriate for most questions. We also **recommend** that, if possible, the comments are written and made available to the external examiners at the time they are considering scripts.

2.4 Procedures in the examination halls

There were a number of annoying glitches at Mill Lane, such as students not being permitted to take the question papers away, and examiners having difficulty finding the space or the tools to write corrections on whiteboards. Dr Evans will compile a list of issues and discuss the matter with the Student Registry. We **recommend** that the Undergraduate Office includes a copy of the maths-specific invigilators' script with the Examiners' Memorandum, so that examiners know what the students should be told.

2.5 CATAM plagiarism

We are delighted that there were again no cases of plagiarism in CATAM last year. We believe that the robust line taken by the Faculty, and supported by the University and external examiners is acting as a deterrent, and hope that the University will continue to support a robust line.

2.6 Timing of corrections

The IB examiners announced a correction about 15 minutes before the end of the examination; the Part II examiners chose not to announce a clarification 5 minutes before the end. We thought that both decisions were correct. We did not agree with the IB examiners' recommendation that any correction in the last 30 minutes of the exam should not be announced, because withholding a known correction is likely to increase the risk of student grievance and complaint. The timings of corrections should be noted, and markers should bear in mind that corrections will take longer to reach candidates in colleges.

2.7 Administrative and Computer Officer support

The examiners' reports remarked on the excellent support given by Ms Glory Dalton and Mr John Sutton, in particular, and the administrative staff more generally. We **recommend** that the Chair of the Faculty Board expresses the appreciation of the Board for their hard work.

3 Part IA

3.1 Extra attempts

Last year's change to the rubric led, as expected, to an increase in the number of extra attempts at long questions from 116 to 188. Compared to the total of 4385 attempts, this is not currently a significant addition to the marking load. We do not agree with the suggestion that students who have 6 long answers should have to choose which 5 to submit. We **recommend** that the new rubric should continue to be used without alteration until further notice.

3.2 Difficulty of questions

The success rates for alphas on the long questions on Groups and Statistics were well outside the Faculty recommended range of 40–45%. We again ask all this year's examiners to pay careful attention to previous years' statistics and the guidelines on question setting in the letter from the Chair of the Faculty Board.

3.3 Draft questions on secure server

We are not persuaded that draft examination questions should be exchanged between examiners and the Undergraduate Office using a secure server rather than hard copy.

3.4 Reporting partial marks

The Chair asks for guidance on whether partial marks (e.g. for physics practicals, or for CATAM) for withdrawn candidates who did not sit the written papers should be reported on the Faculty List. The University guide to examiners states unambiguously that they should be. This is indeed our current practice; such candidates are listed at the end of the list as 'Not Classed' (which is distinct from 'Failed').

4 Part IB

4.1 Difficulty of questions

Prof. Anderson and Prof. MacKay (externals) note that some variations in difficulty between questions on different courses is inevitable, but ask whether it is the case that some courses consistently over the years have easier or harder questions. This is one of the issues that the Teaching Committee looks out for, and the answer in Part IB is no. Examiners are instructed by the Faculty to make every effort to aim for a common level of difficulty, which is specified by target alpha and beta success rates, and to use statistics from recent years to judge the desired direction of travel. For example, as the external examiners note, the questions on some pure and applicable courses were found rather difficult last year, but were found easier this year after the examiners made appropriate adjustments. We agree with Prof. Anderson that there must be a recurrent effort every year, in the interests of fairness for that cohort.

We have an impression that the variability in the difficulty of the applicable questions appears to have been rather greater (in both directions) over a number of years than that of the pure and applied questions. We wondered whether this is because the applicable examiner can be more isolated so that there is less

critical discussion of their questions. If so, we **recommend** that the other examiners try to pay more attention to the difficulty of the applicable questions. We also **recommend** that the Director of the Stats Lab consider whether it might be advantageous to create, as a formal role, a second ‘assistant’ examiner (perhaps also for IA Probability) to assist in setting questions, to provide oversight of difficulty, to attend some of the examiners’ meetings, but not to do any marking. We would welcome constructive suggestions from this year’s examiners.

4.2 Class boundaries

We are a little surprised by Prof. Anderson’s comments regarding the apparent uncertainty over the 3rd/Fail and 2.2/3rd boundaries. While typically about half of the examiners are new each year, the other half are there to give continuity. Moreover, the present classification criteria have been in place for quite a while, are commented on each year by the Teaching Committee and discussed by Faculty Board. Most examiners should be aware of the intention of Faculty Board by now, and abide by that rather than finding their own maverick interpretations.

We reiterate some of the classification criteria. There should be no distinction between marks obtained on the CATAM projects and marks obtained on other courses. Examiners only *may* (not *will*) take into account whether candidates have obtained most (presumably more than half) of their marks on only one or two (and no more!) courses (which certainly includes CATAM in Part IB, where the projects are not tied to other courses) – this condition therefore does not apply to three of those who failed. The primary classification criterion is the merit mark, and there should be no separate consideration of the number of alphas and betas. In our recommendations last year, we clarified the possible sorts of secondary criteria, and made it clear that only in exceptional circumstances should there be a significant departure from the merit-mark order. (We have no objection to the small transposition that the examiners did decide on.) There is no mention at all in the classification criteria regarding the candidates’ prospects in Part II, and there is no requirement of competence in more than one or two courses (including CATAM).

In the context of these criteria, it seems to us that the only notable ambiguity is what merit mark is sufficient for a 3rd (or a 2.2), which can sensibly vary a little from year to year if there is variation in the difficulty of the exam. The borderline for this year’s examination (where three candidates with merit mark in the range 200–250 were failed) is rather high in comparison with recent years, and the examiners’ report gives no insight into their reasoning. Hence we **recommend** that examiners are provided with a printout of the bottom of the class list for the last 5 years as a way of providing comparisons and guidance. We **recommend** that examiners are encouraged to focus on the classification criteria, as printed in the Schedules, and as amplified in the letter from the Chair of Faculty Board, and not on the flexible verbal descriptors or anything else. We **recommend** that examiners minute all decisions pertaining to individual candidates, and their reasons, in conformity with university guidelines (https://www.student-registry.admin.cam.ac.uk/files/examiners_meetings.pdf).

As last year, we are reluctant to add too much prescriptive detail to the Schedules in case this ties the examiners’ hands in unusual situations and prevents them from exercising appropriate judgement. On the other hand, unless the situation is very unusual, we urge examiners to give the highest priority to the criteria as set out in the Schedules.

According to our student representatives, students pay no attention to the verbal descriptors and understand that they should simply pay attention to maximising their merit mark (which is the correct strategy), though occasionally, ill-advised by urban myths, some still worry about numbers of alphas and betas (which is incorrect, except insofar as they contribute to the merit mark). We are not therefore worried about trying to fine-tune the verbal descriptors. We agree with Prof. Anderson that the criteria and descriptions should be kept under review.

4.3 Achievement at lower boundaries

We agree with Prof MacKay that it is important to give opportunities for weaker students to show some level of achievement. It is for this reason that we have been exhorting examiners to keep the short questions genuinely short and straightforward, and we are pleased to see year-on-year improvements in this regard (see §2.2). We also note that the Computational Projects in Parts IB and II are intended as opportunities for candidates to show achievement in ways other than on examination questions, and that the classification criteria explicitly say that marks achieved on CATAM should be regarded in the same way as marks achieved in the examination. Our own experience as external examiners suggests that imposition of credit frameworks is as likely to distort and lower standards of assessment, as it is to raise standards of attainment.

4.4 Model answers and marking conventions

We were unconvinced that solutions with crossings-out are unlike what a good student would produce under exam conditions. Nevertheless, we emphasise the model solutions should be legible, and that the intended solution and mark scheme should be clear to any examiner or supervisor who looks at them. We agree with Prof. MacKay (external) that there should be some indication on scripts where marks are lost (e.g. a ring, underline, \times , \neq , or $\not\neq$) and we note that the instructions on script marking already specify that this should be the case.

4.5 Presentation of CATAM

Prof. MacKay (external) suggests that the proportion of marks for presentational or mathematical clarity in CATAM reports, be increased from the present 10% to perhaps 30–40% for a range of report-writing skills. We had mixed feelings. On the one hand, we agree that these are transferable skills, and worth recognition and encouragement; on the other, we worried that assessment would be seen as subjective, and that students already express concern that they don't know what the "right" answers are for full marks. We referred the matter to the Computational Projects Assessment Committee, since its members have relevant experience of assessing projects.

5 Part II

5.1 Difficulty of questions

Last year, we noted that more of the lowest scoring courses, as revealed by the bar charts of marks *after adjustment* for student effect, were Applied and more of the highest scoring courses were Pure. We recommended that this year's examiners make every effort to reduce differences in difficulty between questions on various courses, and that they inform the relevant lecturers of last year's issues and the desired direction of travel towards the middle of the range. The letter from the Chair of the Faculty Board to the 2017 examiners was strengthened to emphasise this point. We are therefore dismayed that the same problem has arisen again.

While some lecturers clearly did make appropriate adjustments, four of the nine highest-scoring courses in 2016 got even higher scores in 2017 when they should, instead, have been moving back towards the middle of the distribution. Comparing the bar charts before and after adjustment for student effect, we think very little of the differences between courses can be written off as due to differences in the quality of the students taking that course; the dominant effect is due to the differences in the difficulty of the questions being set.

The examiners' reports detail the extensive efforts they went to look at the range of questions attempted by each candidate in the light of the statistics, with the result that they moved the borderlines lower down the list so that borderline Applied candidates were treated generously and not disadvantaged by their course selection. Nevertheless, we **recommend** that much greater efforts are made towards homogenization of difficulty. In particular, we **recommend** the following measures:

- Writing specifically to the lecturers of the courses that were significantly out-of-line with the Faculty guidelines in the previous year.
- Requiring all lecturers to sign a sheet saying they have considered the statistics from previous years when handing in their questions to the Undergraduate Office.
- Asking lecturers to aim for the Faculty's target alpha rate rather than assuming that any discrepancies are due to differences in the quality of the students.
- Heads of Departments should try to ensure that the slate of examiners has significant expertise across the full range of courses.
- Examiners and checkers need to take more responsibility, and challenge lecturers if they think that the questions are not appropriate.
- Examiners are reminded of their freedom to tweak mark schemes in the light of the answers in a way that distinguishes more clearly between good and adequate understanding.
- New lecturers to Cambridge should be given explicit guidance and mentoring regarding the importance, in the Mathematical Tripos, of setting questions at the appropriate level of difficulty.

5.2 Model answers

Prof. Sobolev (external examiner) commented that most of the model solutions he was sent were illegible, and some solutions were incomplete. We weren't persuaded by his suggested remedy of having solutions typed in LaTeX – we believe that a hand-written solution showing what a good student would write gives a much better indication of the length of the question. We totally agree with him that solutions should be legible and kept up-to-date with any changes in the question.

The lecturer is responsible for providing the *first* drafts of questions and solutions, and for making every effort, as the specialist, to use previous statistics to provide questions set at the appropriate level of difficulty. The examiners are unequivocally responsible for the questions on the examination paper, for comparing questions on different courses, and for making sure that the solutions and mark schemes are accurate and clear to any examiner or supervisor who looks at them. If necessary, solutions should be written out again by the examiner. While the preparation timetable is often tight, we **recommend** that the Chair encourages examiners to exercise quality control on what is sent to the external examiners.

5.3 Mark sheets

We thought the suggestion of adding the average mark for a question beside each individual's mark for that question would probably clutter the mark sheet and obscure the information it currently presents. The statistics on questions are, of course, routinely provided to examiners in a separate document.

6 Summary of recommendations

(The exact recommendation is described in the section indicated.)

- 2.1 The Chair to impose typesetting conventions between examiners' meetings.
- 2.3 Expectation of the content of the comments on questions. Available to externals.
- 2.4 Inclusion of invigilators' script with examiners' memorandum.
- 2.7 Thank Ms Glory Dalton and Mr John Sutton for their efforts last year.
- 3.1 Continued use of the new rubric.
- 4.1 Increase scrutiny of the applicable questions in IA and IB.
- 4.2 Provision of history of decisions at the bottom. Focus on the classification criteria as printed. Minuting of all decisions.
- 5.1 Various measures to improve homogenization of difficulty.
- 5.2 Greater effort to improve the quality of solutions sent to external examiners.

Nilanjana Datta	Oscar Donlan	Henry Wilton	Berry Groisman
Lorna Ayton	George Robinson	Mike Tehranchi	András Zsák
John Lister (Chair)			

November 10, 2017