

Teaching Committee report to the Faculty Board

Examinations for Parts IA, IB and II of the Tripos, 2012

1 Introduction

The Committee met three times, for two hours each time, to consider the examinations. We looked, for each part of the Tripos, at:

- the Examiners' reports;
- the External Examiners' reports;
- the Examiners' comments on their questions;
- the examination statistics (take-up rates, etc, this year and last);
- 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);
- the regression analysis of Part II question marks;
- the Schedules booklet

We also looked at letters sent each year by the Chair of the Faculty Board to the Examiners (individually), the 'Interpretation of Marks' letter sent by the Chair of the Faculty Board to Directors of Studies and the letter sent to examiners shortly before the examinations explaining the arrangements for the examination ('Examiners' Memorandum').

We note with pleasure that 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:

I am pleased to find in place a rigorous and efficient procedure for setting, checking and reviewing examination papers. Marking of examination scripts was conducted fairly, with great attention and care. It is obvious that the Mathematics Department has well tested and finely tuned examination mechanism. (Part IA)

My conclusion about the exams process is that it is rigorous, and has been conducted with great care by a thorough and committed group of examiners. The great majority of candidates showed considerable achievement, and those at the top end showed impressive ability. I am satisfied with the outcome. (Part IB)

Exam papers were produced to a very high standard. The system of short questions for relatively routine material and longer questions requiring in-depth understanding seems to work well. Completed CATAM projects that I looked at were of a very high standard. (Part IB)

The breadth of topics covered in the exam questions is impressive, and possibly unique not just nationally but internationally for an examination at this level. The exam performance of students at the very top end is spectacular, and I note that many students at the lower end of the degree classification still demonstrate good mathematical maturity and knowledge. (Part II)

The standards in general are high, and at the top end some candidates give sensational performances. Even at the bottom end of the range candidates are able to demonstrate some achievement, with all scripts showing some evidence of quality learning, even if the quantity is minimal. Thus the standards of the examination appear to be entirely appropriate for the group of students who take it. (Part II)

Standards in the Mathematical Tripos are high, comparable with or above those at other leading UK institutions of which I have direct knowledge. The best candidates demonstrate great ability over a wide range of subjects, but even lower on the class list there is strong evidence of mathematical maturity and skill. The structure of the examination ensures that it serves as a useful discriminant across the ability range of the candidates. (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 to pick up from the previous year's Examiners' reports. We are pleased that the annual meeting of the Directors of Studies has again requested a copy of this report in time to be able to contribute to the Faculty Board discussion.

2 General Matters

2.1 Errors

This year the following errors were reported, out of a total of about 280 questions: one error in Part IA; three errors in Part IB; and six errors in Part II; any error is unfortunate, but this year there were significantly fewer errors than occurred in many recent years. In Part II, two issues arose after the examination, one concerning a convention of which not all candidates were aware and the other concerned a possibly misleading hint.

2.2 Differences in difficulty between questions

There was considerable disparity in the levels of difficulty of questions and courses, both as perceived by candidates (online questionnaire) and as measured by the examination statistics (take up rates, average scores and percentage of alphas), in all parts of the Tripos this year. This was alluded to by the Part IA examiners, the Part IB examiners, and one of the Part IB external examiners (the other remarked that such differences would only be problematic if there was a long-term trend). A large number of students expressed annoyance and disappointment at the unfairness that resulted from this disparity in the levels of difficulty of the questions.

One might have expected a given course to have harder and easier questions, giving a balanced overall package and, if this did not happen, one might have expected a given subject area to have hard and easy courses giving overall a balanced package. However, the alpha rate and average mark for long questions in the three different areas in the table below shows that there was essentially no balancing in Parts IA and IB:

2012	Part IA			Part IB		
	Take up	Average Mark	Alpha rate	Take up	Average mark	Alpha rate
Pure	51.8%	12.9	0.50	38.0%	12.2	0.37
Applicable	50.5%	11.0	0.21	24.0%	10.2	0.29
Applied	61.2%	15.4	0.67	42.5%	15.6	0.69

Last year, the picture was similarly disparate, with Pure and Applied essentially reversed:

2011	Part IA			Part IB		
	Take up	Average Mark	Alpha rate	Take up	Average mark	Alpha rate
Pure	58.3%	13.9	0.52	43.7%	15.4	0.65
Applicable	53.6%	15.2	0.61	24.7%	15.2	0.67
Applied	52.7%	12.6	0.44	38.6%	13.6	0.51

The disparity between individual courses was even more stark. At the ends of the range the alpha rates in 2012 (with corresponding alpha rates for the same courses in 2011) were as follows.

Part IA: 0.21 (0.61) for Probability; and 0.74 (0.42) for Vector Calculus.

Part IB: 0.20 (0.77) for Geometry; and 0.78 (0.64) for Methods.

We believe that the problem arises for two reasons:

- there is no guidance to examiners and lecturers to indicate how hard a question should be (except in the case of short questions);
- examiners are reluctant to take collective responsibility for the level of difficulty of each question.

Of course, it can be difficult to gauge the difficulty of an individual question, because it is affected by a range of factors including the difficulty of the whole course, the success of the lecturer, the extent to which similar material was covered in lectures or on examples sheets, etc, though one can often obtain good guidance from the exam statistics relating to similar questions in previous years.

Our experience on examination boards is that individual examiners may have strong views about the desired level of difficulty and may be reluctant to change the level of difficulty of their questions even when urged to do so by other examiners.

We believe that a target average mark or average alpha rate for each question or course, agreed by the Faculty Board, could be effective in smoothing out levels of difficulty. We accept that candidates will find some questions harder or easier than had been anticipated by the examiners, but believe that, with an agreed target and guidance from previous years' statistics, the recent large variations can be reduced.

We therefore **recommend** that the Faculty Board issues guidance on this matter, and suggest that targets of (40 – 45)% alpha rate on long questions and (55 – 60)% beta rate on short questions would be appropriate. These proposals for targets reflect the historical data for the long questions and for the short questions reflect our view that the historical rates are a bit low and that the short questions would better serve their purpose if a higher rate were aimed for. ¹

2.3 Invigilation

There were again this year problems with invigilation: in one room (for Part IA, Paper 2), no invigilator was present at the start of the examination and an examiner had to stand in. The Part IA Examiners also note that some of the invigilators were not sure what they were supposed to do (for example with the cover sheets). We **recommend** that the Faculty Board draws these matters to the Board of Examinations.

Two other issues arose:

1. For Part IA, the examiners report that a substantial number of candidates wrote some or all of their answers on the wrong side of the paper (with the punched hole at the top right corner), causing examiners and checkers considerable inconvenience. They surmise that the paper was not laid out carefully. We **recommend** that the Faculty Board asks the Board of Examinations to include the matter of careful layout of paper in the briefing to invigilators or others responsible. We also **recommend** a change to the e-mail sent to Part IA candidates before the examination, so that it refers not just to writing on one side of the paper, but to writing on the side with the hole in the top left corner.

¹For Part IA for the years 2003 to 2012, the overall alpha rates are 0.41, 0.45, 0.44, 0.48, 0.38, 0.45, 0.40, 0.53, 0.49, 0.56. For Part IB, the corresponding figures are 0.39, 0.46, 0.35, 0.39, 0.47, 0.40, 0.38, 0.47, 0.59, 0.53. For Part II 2005 (before which the Tripos was split into Alternatives A and B) to 2012, the figures are 0.46, 0.52, 0.46, 0.49, 0.45, 0.52, 0.51, 0.48. The high alpha rate in some parts of the Tripos in recent years is a result of a few very easy and unchallenging questions.

For Part IA for the years 2007 - 2012, the beta rates for short questions are 0.56, 0.44, 0.38, 0.48, 0.53, 0.57. For part IB, the corresponding figures are 0.55, 0.44, 0.43, 0.49, 0.57, 0.46 and for Part II they are 0.49, 0.43, 0.54, 0.50, 0.54. These beta rates suggest to us that the shorter questions are not quite as straightforward as the Faculty Board intends, which why we are proposing a beta rate higher than the historical average.

2. We decided that it would be helpful, in relation to the above matter but also to other matters (for example, to announce the presence of an examiner in the building for the entire three hours of the examination as stated in the Memorandum to Examiners² and not just the first 30 minutes) if the invigilator were to read out a maths-specific script at the beginning of the examination, and we **recommend** that the Faculty Board ask the Board of Examinations whether this can be put in place for this year.

2.4 The merit mark dogleg

Dr Hammerton (Part IB External) had not found it easy to understand from the single order of merit list the method of ordering and wonders whether two lists would be clearer. However, the Examiners' Report states that the current system (with marginal flags) implemented this year was very helpful. It seemed to us that there are significant dangers in having two lists and that it would be better for the Chairs of Examiners to explain carefully to new External Examiners the method of ordering than to make yet another change.

2.5 Faculty Secretary and Computer Officer

All the Examiners' reports remarked on the excellent support given by the staff of the Faculty Office, and in particular by Amy Dittrich. We **recommend** that the Chair of the Faculty Board expresses the appreciation of the Board for their work.

The Examiners in all Parts of the Tripos found the expertise of Mike Rose, Chris Mortimer and John Sutton invaluable. We **recommend** that the Chair of the Faculty Board expresses the appreciation of the Board for their work.

3 Part IA

3.1 Question style

The External Examiner, Professor Borovik, would like to see an increase in the number of questions that examine 'unseen' material. It is mentioned in the Examiners' report that three questions had an average mark of 18 or above and a further 10 had an average mark of 15 or above. The overall alpha rate this year was 0.56, which is the highest for 10 years and much higher than in many of these years. It seemed to us that it was this unusually high number of unchallenging questions was largely what lay behind Professor Borovik's concerns and that in a more normal year the balance is in fact about right. The recommendation in section 2.2 above would, if adopted, go some way towards alleviating these concerns.

4 Part IB

4.1 Relation between marks and quality marks

Professor Greenlees (External Examiner) proposes removing the direct link between marks and merit marks on the grounds that examiners sometimes ignore their mark scheme in the interests of justice. We discussed this idea (this year, as well as last year), but concluded that it would add additional variables (and more subjectivity) to the process with little benefit; in most situations, a well considered mark scheme, constructed with the alpha and beta points in mind, should suffice. We wondered whether the External Examiners see the Teaching Committee report; it seems not, from Professor Greenlees's remarks. We **recommend** that our report (as well as the Faculty Board's response to the General Board) be sent to External Examiners.

²It could be stated more clearly here.

4.2 Tracking of scripts and individual questions

Professor Greenlees (External Examiner) wonders whether it is time to use modern technology to replace our (recently introduced) ‘early technology’ of scanning cover sheets. He suggests using bar codes for each question. We felt that the current scheme works well enough and that, since the current scheme has only quite recently bedded down, we should not consider a radical change yet.

4.3 Guidance at the lowest borderline

Professor Greenlees (External Examiner) suggests that it might be useful for the Faculty Board to give further guidance on how to treat the candidates who would previously have been awarded an Ordinary. However, Faculty Board guidance on this point has been unambiguous: they should fail. It is possible that the Part IB examiners were not aware of this guidance this year, but we felt that no action need be taken, especially if this report is sent to the Chairs of Examiners (if it is not, then we **recommend** that in future it is).

4.4 Marking style

Professor Greenlees (External Examiner) would have found it useful if examiners had adopted a uniform style of marking so that, for example, it was clearly indicated how many marks were lost (or gained) at each point (in the form 3/5 for example). We felt that this would be an additional burden on very hard-pressed examiners; the same information would be available from the marking scheme. However, we agreed with Professor Greenlees that the final marking scheme should be made available to the External Examiners; this is already clearly stated in the current instructions to examiners (in the letter from the Chair of the Faculty Board to Examiners).

4.5 CATAM

We noted from the CATAM Director’s report that an anomaly was detected in the statistics pertaining to one of the questions, and that all attempts at this question were accordingly re-marked. We were very impressed with the process that led to the complete rectification of an unfairness which might easily have been missed altogether and we **recommend** that the Chair of the Faculty Board personally commends the CATAM Director for his diligence and attention to detail.

4.6 CATAM plagiarism

We were surprised to read that no action (apparently) was taken against two candidates who appeared to have colluded. Our understanding of the University policy (there is a flow chart) is that if there is evidence of plagiarism but the candidates concerned do not admit to it, the case should be referred to the University Advocate. We **recommend** that the Director of CATAM be asked to make a brief statement to the Faculty Board clarifying this matter.

5 Part II

5.1 Late questions

The examiners report that the questions proposed by some lecturers were significantly late, two weeks late not being uncommon and three weeks late not being unknown. This seems to us to be a recipe for producing an unsatisfactory and error-strewn examination. The time-table between receipt of questions and transmission of questions approved by the relevant examiners and checkers to the external examiners is very tight and any slippage will result in questions being sent out that are not in an appropriate form for the external examiners to scrutinise. We **recommend** that the Faculty Board devises a procedure to ameliorate this situation.

Currently, the Faculty Office sends a reminder a week before the deadline for the lectures to provide drafts (which is itself a week before the start of Lent full term) and several later reminders for the miscrants. We suggest instead a series of reminders of increasing severity culminating in a report to the Faculty Board (as for Part III). For example, immediately after the deadline, the Chair of Examiners could send a reminder and a week after the deadline the Chair of the Faculty Board could send a reminder with the threat of a report to the Faculty Board if the questions are not provided at once.

5.2 Errors

Professor Schroers (External Examiner) suggests appointing proof readers who have not already been involved in the setting of the questions to reduce the number of errors. We noted that there is already a checker assigned to each course and we thought that a further stage of checking, while possibly reducing the number of errors, would add an extra layer to the examining procedure (with implications for both security and timing). There are sufficient examiners to ensure that there are almost no errors if they check their own and others' questions with care; and indeed there were only 6 errors in total (in around 170 questions) this year in Part II, and nearly all were very trivial (for example, $T \ll k_B E_F$ instead of $k_B T \ll E_F$ would have been spotted very easily by any candidate who had attended the course).

5.3 Alpha weighting

Professor Schroers (External Examiner) expressed concern about the weight given to alphas; he suggests a smoother formula. This matter is also raised by Professor Wilson (External Examiner). The Faculty Board has discussed this very many times in the last 10 years and we do not recommend re-opening the issue yet again. We note, though, that last time it was discussed in depth various different models were suggested and tested on existing class-lists and there was very little difference between the ordering of candidates.

It should be borne in mind that both the candidates and the examiners are now well used to and understand the existing system of quality marks.

5.4 Range of skills

Professor Schroers (External Examiner) comments on the narrow range of skills tested by the Tripos: basically just the ability to do examination questions. We were surprised by his comment that few questions require genuine problem solving skills or test understanding of the significance or context of a mathematical results, but we appreciated the general point. He suggests a paper with essay-style questions (similar to the old Paper 4), student presentations and extended essays. We noted that some of these ideas are built into the course at Oxford, though in a very limited way.³

These matters are currently under discussion by a Faculty Working Group, so no further action is required at present.

5.5 Examination structure

We were pleased to see Professor Wilson's (External Examiner) endorsement of our non-modular examination structure and also Professor Young's (External Examiner) view that the structure of the examination (with Section I and II) ensures that the examination serves as an effective discriminant across the ability range of candidates.

5.6 Examiners' fees

We hope that Professor Wilson's complaint regarding the inadequacy of the fees paid to External Examiners will be considered by the appropriate body.

³Two members of the Committee had been External Examiners at Oxford last year and did not advocate the introduction of similar schemes in the Mathematical Tripos.

5.7 CATAM marks

We noted from one of the External Examiners' reports (Professor Young) that there was a problem that arose after the final meeting of Examiners (and after the class list was released) over incorporation of marks from the computational projects into the marks list. We understand that this did not affect the classification. We expect that next year's examiners will be vigilant in ensuring that this problem does not recur.

5.8 Efficacy of C-courses

In trawling through the examination statistics, we were discouraged (again) by the low take-up rates of some of the C-courses, as shown in the following table.

<u>Course</u>	<u>2012 take-up rate (%)</u>	<u>2011 take-up rate (%)</u>
Coding and Cryptography	17.9	24.1
Groups and Geometry	3.0	6.8
Number Theory	45.0	48.4
Topics in Analysis	5.7	7.7
Statistical Modelling	4.8	6.8
Classical Dynamics	12.4	15.1
Cosmology	6.4	5.2
Dynamical Systems	8.1	8.6
Further Complex Methods	12.8	15.8
Mathematical Biology	6.0	8.1

Of course, we cannot expect that all 10 courses will have the same take-up rate as Number Theory; nevertheless, it cannot be right that so many of these courses, which are intended for the less strong students, have such low take-up rates.

We gleaned from the student questionnaires some reasons for these low take-up rates. For Cosmology and Dynamical Systems there is just too much material in the schedule.⁴ For Groups and Geometry and Topics in Analysis it is hard to escape the conclusion that it is the courses themselves that are unsatisfactory.

We also noted the exceptionally low take-up rates for Partial Differential Equations (1.8% in 2012; 0.6% in 2011). It seemed to us that a review is overdue. For example, Dynamical Systems could be designated a D-course (with no change of schedule) and be replaced by something akin to the old and much liked Partial Differential and Integral Equations course, while the current Partial Differential Equations course is found a berth in Part III.

We **recommend** that the Faculty Board reviews these courses urgently.

Clive Newstead	Jack Williams	Colm Caulfield	David Mestel
Susan Pitts	Berry Groisman	Imre Leader	
John Lister	Tom Fisher	Stephen Siklos (chair)	

November 7, 2012

⁴Recall that when these courses were first mooted, it was suggested that they might have the same amount of material as a 16-lecture course but be lectured at a more leisurely pace with more examples.