

Teaching Committee report to the Faculty Board

Examinations for Parts IA, IB and II of the Tripos, 2018

1 Introduction

The Committee met four times, for two hours each, to consider the 2018 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 academic standards set by the examination were very high. The calibre of the best students means that the exams have to have substantial, hard questions while at the same time including questions accessible to weaker students. (Part IA)

The standards of the examinations and of the qualifications are very high. The processes for assessment, examination and the determination of awards are sound and fairly conducted. (Part IB)

The examiners take a great deal of care in both the setting and marking of examination scripts and the Faculty of Mathematics has robust systems for checking and compiling marks. (Part IB)

The standard of the examination papers was very high, and the performance of the candidates was typically exceptionally strong. In a period of grade inflation across the sector, it is reassuring that Cambridge continues to hold strong and not devalue its own currency. (Part II)

I was extremely impressed with the quality of the students, and with the rigour of the examination process. Part II Maths continues to maintain the highest standards in terms of both the breadth and depth of its courses, and to provide a worthy and challenging hurdle for the UKs strongest students to compare themselves against. (Part II)

The level and breadth of Pure Mathematics courses offered was excellent, and the standard of the exam questions was high. (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 nine errors were discovered before, during or after the examinations, in a total of about 280 questions: no errors in Part IA; two errors in Part IB; seven errors in Part II.

The examiners followed last year's recommendation that typesetting conventions should be applied before the second or third examiners' meeting, if necessary by the Chair. They comment that this worked well, allowing the examiners to focus on the mathematical content and difficulty in the meeting, rather than being distracted by issues of style. We strongly **recommend** that this practice continue.

2.2 Reuse of old questions and model answers

The examiners in Part IB were notified by some concerned students after Paper 2 that all the questions in Quantum Mechanics and Variational Principles were exact duplicates of questions from previous years (2007 & 2011). Anticipation that this might recur in Papers 3 and 4 was shared via social media, particularly in some larger colleges, and concern was expressed that this might disadvantage those students not in the know. The questions on Papers 3 and 4 were indeed duplicates from previous years, but not years that might easily have been anticipated. There is no evidence, statistically speaking, that any students were disadvantaged. Nevertheless, this incident should serve as a wake-up call to any examiners thinking it is safe in an internet age to recycle multiple questions.

We **recommend** that: (i) examiners do not reuse any question from the previous at least 6–8 years (ii) examiners be aware that it is very easy to identify reused questions on early papers from the online resources available (iii) examiners should notify the committee at the second meeting of any prior instances of the questions they are proposing to set (iv) the majority, preferably all, of the questions on any course should be new, though appropriateness should be given much greater priority than novelty. The guidance for examiners will need updating.

A number of student complaints in the end-of-year questionnaires alerted us to the fact that a complete set of LaTeXed answers to IA Probability questions 2001–2017 was circulated by social media shortly before the Part IA examinations. The answers were not copies of those held in the undergraduate office, and were presumably prepared by a supervisor and given to their supervisees. There are issues with fairness and relative advantage within the student body. We **recommend** that the Faculty reiterates its policy that undergraduates should not be given model answers to questions, beyond the usual going through recent questions in supervisions, and perhaps providing solutions to questions the students themselves have attempted. For students to work to gain an advantage by doing more questions under their own steam is one thing, but we strongly deprecate, on educational grounds, the idea that students should be prepared for exams by being fed a set of answers to learn.

2.3 Administrative and Computer Officer support

The examiners' reports remarked on the excellent support given by Ms Kati Sexton, Mr Mycroft Rosca-Mead, Ms Gaynor Lindsay-Walters 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 Difficulty of examination

This year's examination was found significantly more difficult than that in recent years, with marks about 20% down at each border and the number of questions attempted down by about 5%. In IA an unexpectedly hard exam is, of course, the same for everybody (except the physicists), but

some may find it significantly more disconcerting than others. We **recommend** that the IA exam briefing include the advice ‘expect the unexpected’. Though marks were down across the board, the criterion for a 3rd was similar to last year ($2\alpha + \beta > 9$), with the result that there were 9 failures. While we agree it is not appropriate to allow candidates to progress who are unequipped to do so, we are somewhat surprised by where the line was drawn.

While the overall 42% alpha rate on long questions was within the target range of 40–45%, the 49% beta rate on short questions fell well short of the target 65–70%. Prof. Watts (external examiner) comments if the short questions were more straightforward (easier) then it would simplify matters at the lower end of the scale while barely affecting the better students. We agree that short questions are very important as a means for giving even the weakest students opportunity to demonstrate success in learning.

We **recommend** that this year’s examiners pay careful attention to the statistics from previous years and the guidelines on question setting in the letter from the Chair of the Faculty Board.

3.2 Draft questions on secure server

We are still not persuaded that draft examination questions should be exchanged between examiners and the Undergraduate Office using a secure server rather than hard copy. Any breach of security, perhaps thorough human error, would be catastrophic.

4 Part IB

4.1 Typographical conventions

There seems to be some confusion about typographical conventions, with some examiners remembering that, once upon a time, (a), (b), (c) was used to denote logically distinct parts of questions and (i), (ii), (iii) was used for related parts of a question. Precisely because students, and more recent examiners, were completely unaware of this obscure historical practice, this convention was formally abandoned in 2015, and the current conventions for labelling parts are those that are circulated to examiners.

4.2 Queries during the examination

The examiners’ report notes that ‘there seemed to be an excessive number of queries requiring no action this year and some instances of students wanting to engage in discussion about their queries’. The Teaching Committee was very much against suggestions that queries should be written, and delivered via an examination administrator, noting that it is vital that real mistakes are identified and disseminated easily and accurately. At the same time, we encourage examiners to give a simple ‘the question is correct as written’ response to any query that has not identified either a mistake or a genuine ambiguity that **requires** clarification to all students.

4.3 Difficulty of questions

There was good balance in the difficulty of Section II questions, but the Section I questions on pure and statistics were on average not as accessible as they might have been. We **recommend** that next year’s examiners again pay attention to the need to keep Section I questions accessible.

Last year we noted 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. This year, with a very experienced examiner, the variability was less, but we remain concerned that the applicable examiner is potentially more isolated with less critical discussion of their questions. We again **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, and to attend some of the examiners’ meetings, but not to do any marking.

4.4 Number of questions

The examiners suggest that Faculty Board might consider reducing the number of questions per course (currently 4 long and 3 short for a 24 lecture course; 3 long and 2 short for a 16 lecture course; 2 long and 2 short for a 12 lecture course). The Teaching Committee had a lot of sympathy with this view, noting that there are more questions per course than in either Part IA or Part II, and that fewer questions might encourage students to take a slightly broader range of courses. It was not obvious to us how to reduce the number of questions while maintaining an appropriate balance in both short and long questions between courses of different lengths. However, bearing in mind the different purposes of long and short questions, we **recommend** that the number of short questions on a 24 lecture course be reduced to 2. All courses will then have two short questions, while the number of long questions (4, 3 and 2 respectively) continues to be reflective of the course length.

4.5 Borderline processes and question structure

Prof. Anderson (external examiner) notes that while students do know the merit mark formula, the approximate borderlines from previous years, the percentage guidelines and the classification descriptors, they may not know the precise process by which examination boards decide borderlines. The student representatives on the Teaching Committee were completely satisfied that they already knew more than enough to guide their examination strategy.

Prof. Harlen (external examiner) welcomed the fact that classification is not entirely algorithmic and leaves space for some academic discretion and judgement, but was concerned that the basis on which discretion is exercised should be clear. We agree! In 2016 Faculty Board agreed a description of when and how ‘other factors’ might be taken into account, which was communicated to the 2017 examiners, but was omitted by mistake from the information given to the 2018 examiners. The omission has been rectified, and future examiners will receive this guidance. The guidance is very much along the lines recommended by Prof. Harlen.

Prof. Anderson also wondered whether questions should have a uniform structure. In part, this may be prompted by the misunderstanding regarding typographical conventions described above. The Committee was firmly of the view that imposing a uniform structure would be an unnecessary constraint that would have a negative impact on the design of suitable questions for each course.

4.6 Marking conventions

Prof. Harlen suggests denoting partial marks as fractions, e.g. $3/5$, to show where marks are lost, and that examiners should comment explicitly that a solution was or wasn’t worth an alpha when awarding $14/20$ and $15/20$. The Teaching Committee noted that examiners are already instructed to show where marks are lost by means of crosses, underlines, rings, etc. and that student answers are often not in sequence so that the ‘ $3/5$ ’ may be dispersed through the script. Examiners are very aware of the big difference between $14/20$ and $15/20$. We were unpersuaded that there was added value in increasing the amount of writing, and thought it might be a distraction during marking.

4.7 Transferable skills

We note with pleasure the positive comments by Prof. Harlen regarding the high uptake of CATAM, and the opportunity it provides to develop computational and written communicational skills to complement broad mathematical skills. Cambridge supervisions develop oral and written communication skills, while challenging assignments provide opportunities for team work with a supervision partner or friendship group. Colleges and societies also provide opportunities to develop a wide range of social and organisational skills. Our belief, supported by strong employment statistics, is that students are acquiring transferrable skills without need for summative assessment in place of some mathematical content.

4.8 Advice on strategy

Prof. Harlen comments that some candidates made poor tactical choices about how to balance short and long questions. Students are generally advised to optimise their merit mark. First-year students are given more specific advice in a pre-examination briefing at the start of the Easter Term. All students receive an email containing examination advice, in which they are invited to take note of the straightforward nature of short questions when considering their strategy. Individuals should receive tailored advice, appropriate to their abilities, from their Director of Studies. Unfortunately, some candidates may not listen to advice.

5 Part II

5.1 Difficulty of questions

In recent years, there have been persistent differences in the difficulty of questions on various courses, with a handful of Pure courses, in particular, repeatedly having alpha rates well in excess of the target range. A number of new measures were put in place following last year's report on examinations, in particular putting a greater emphasis on the responsibility of lecturers to consider Faculty guidelines regarding difficulty and the desired direction of travel from previous years' statistics. We are very pleased to note that these measures were largely successful, both in terms of the examiners reporting 'more uniformity in difficulty between different areas of the Tripos than in previous years' and in terms of most previously out-of-line courses producing results much closer to Faculty targets.

We **recommend** that all the new measures remain in place. Like the letter to examiners, the letter to lecturers should come from the Chair of the Faculty. We emphatically agree with the examiners' **recommendation** that similar vigilance be shown next year for courses which may have been slightly too easy or too hard this year, and special vigilance shown towards courses that have been systematic outliers over several years.

5.2 Queries during the examination

Despite the best efforts of examiners, the size and complexity of Part II makes it almost inevitable that more errors will slip through into the Part II papers than in Parts IA and IB; this year's figures are quite typical. There will thus be more queries to address in the examination halls in Part II, with the additional complication of possibly needing to contact the responsible examiner. We **recommend** that the Part II examiners consider whether having three duty examiners at each examination might mitigate the stress of a cluster of queries, and facilitate a more rapid response. We note that the Board is larger than in Parts IA and IB.

5.3 Paperwork

Prof. Johnson (external examiner) suggests that it would have been easier to judge the suitability of draft questions, if he had been sent the statistics on questions from previous years. We **recommend** that it should be standard practice to send two years of past statistics with the bundle of draft questions to external examiners in all Parts of the Tripos.

The three external examiners all commented that their consideration of scripts on the Monday was delayed and made less efficient by the unavailability of statistics and box plots until later in the day. In part, this was caused by at least one examiner missing the Thursday deadline for the return of scripts, and the subsequent unavailability of a mark checker. It should be a matter of both professionalism and consideration for others that deadlines are not allowed to slip, and that the external examiners should always have a full set of paperwork available to them at the very start of their day's work. We **recommend** that procedures and instructions to examiners are tightened up. We also **recommend** that when checkers are assigned by the undergraduate office the expectations and timetable should be clearly explained.

5.4 Merit mark

Perhaps associated with the chaotic start on Monday, the external examiners expressed some uncertainty about the ranking order presented to them. The concern is in our view misplaced. Even in the extreme example constructed by Prof. Johnson of a candidate with 15 alphas but few marks otherwise, the possible merit marks would be either $M_1 = 555$ or $M_2 = 450$. The former is not even close to a First and the candidate would therefore be assessed on the basis of M_2 for whether they should get a 2.1 or a 2.2. In practice, there is no ambiguity about which borderline a candidate might be close to, and the marks processing program ensures that candidates are ranked in the appropriate order near each borderline. We **recommend** that chairs and external examiners are provided with more details of the algorithm.

5.5 Equality, diversity and inclusion

Prof. Johnson (external examiner) felt uncomfortable that the examination board was all male. As in 2015, the Teaching Committee did not believe that the gender balance of the examiners was relevant to the performance of the candidates. All candidates are considered anonymously, with their gender unknown. Moreover, previous statistical analyses have failed to show any appreciable difference in the strategies pursued by male and female candidates, in contrast to stereotypical assumptions that are too easily made. The Faculty is fully committed to providing a level playing field with respect to all issues of equality, diversity and inclusion, and has a number of committees and working groups examining detailed statistical evidence related to these issues.

5.6 Model answers

We are pleased that Prof. Sobolev (external examiner) saw some improvement in the quality of the solutions sent to him with draft questions. There is still room for improvement, and we support him in his request for more carefully prepared solutions. As last year, we are not in favour of having solutions typed in LaTeX, since a hand-written solution showing what a good student would write gives a much better indication of the length of the question.

6 Summary of recommendations

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

- 2.1 The Chair to continue imposing typesetting conventions between examiners' meetings.
- 2.2 Restrictions on examiners' use of past questions.
- 2.2 Reiterate restrictions on providing model answers to students.
- 2.3 Thank Ms Kati Sexton, Mr Mycroft Rosca-Mead, Ms Gaynor Lindsay-Walters and Mr John Sutton for their efforts last year.
- 3.1 IA exam briefing for undergraduates to include 'expect the unexpected'.
- 3.1,4.3 Attention to the need to keep short questions short and straightforward.
- 4.3 Possible increase of scrutiny of the applicable questions in IB (and IA).
- 4.4 Reduction in the number of short questions on a 24 lecture course in Part IB from 3 to 2.
- 5.1 The new measures to even difficulty of questions to remain in place, vigilance to be maintained.
- 5.2 Part II examiners to consider whether to have 3 duty examiners at each examination.
- 5.3 Statistics to be sent with the draft questions to external examiners in all Parts. Tightening of procedures for getting everything ready for the external examiners.
- 5.4 Details of the marks-processing algorithm for the chairs and external examiners.

Sergio Ballacado	Oscar Donlan	Alexei Kovalev	Berry Groisman
Julia Gog	Xiao Ma	George Robinson	Mike Tehranchi
Richard Jozsa	András Zsák	Jonathan Evans	John Lister (Chair)

November 8, 2018