

## NST part II, History and Philosophy of Science

### Senior Examiner's Report, 2013

This year 40 candidates sat the HPS Part II Examination, a similar figure to those of the past three years (Table 1). Most chose option A, with dissertation, but 13 (32%) chose option B, without. This is broadly in line with recent years (Chart 2). The final results for the HPS Part II comprise 7 Firsts (18%), 29 Upper Seconds (72%) and 4 Lower Seconds (10%). This represents a smaller proportion of first-class awards than in the past few years (Chart 1) but the examiners felt that the distribution was a fair reflection of this cohort's performance.

Year	First	Upper Second	Lower Second	Third	Deserved Honours	Total	A	B
2013	7	29	4	—	—	40	27	13
2012	16	19	1	—	1	37	23	14
2011	11	25	3	—	—	39	28	11
2010	11	26	2	—	—	39	25	14

Table 1: Distribution of HPS Part II marks, 2010–13

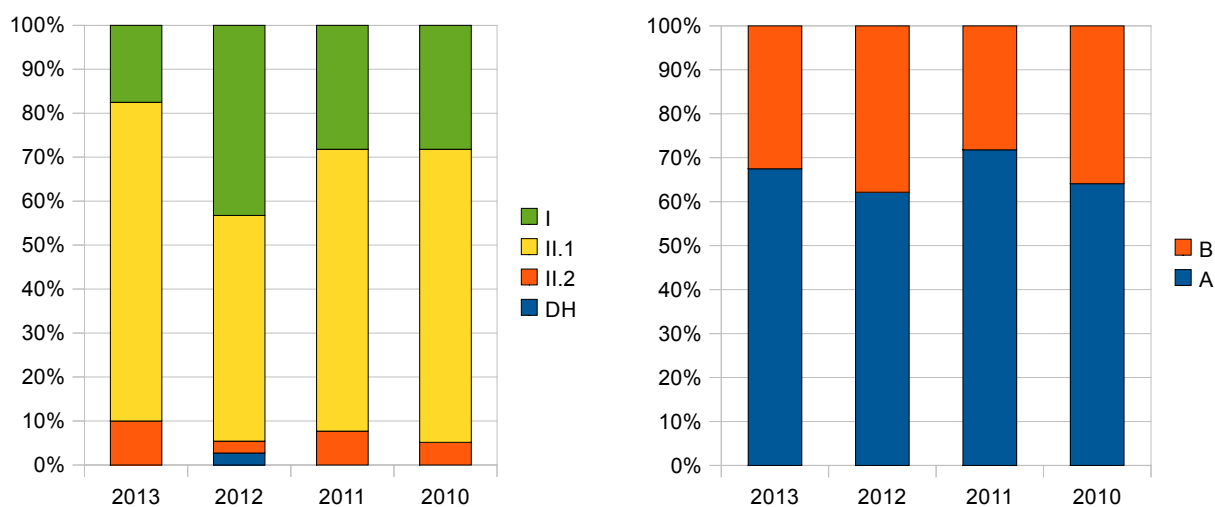


Chart 1 (left): Percentages of HPS Part II classes, 2010–13

Chart 2 (right): Percentages of Option A (dissertation) and Option B (no dissertation) candidates, 2010–13

A further 5 candidates sat Paper 2 (Early Medicine) as BBS Minor Option 65, 3 took Paper 5 (Modern Medicine) as BBS Minor Option 66, much the same numbers in recent years. However, only 22 sat History and Ethics of Medicine as BBS Minor Option 45, a drop of some 50% over the past two years but a similar number to 2010 (Chart 4). The HPS examiners do not formally class these performances but pass on the marks to the BBS Examining Board. Informally, though, I note

that 7 (32%) of HEM students were awarded Firsts, a larger proportion of than in the recent past (Chart 3). 11 (50%) received Upper Seconds, 2 got Lower Seconds, one a Third, while one candidate failed (Table 2).

Year	First	Upper Second	Lower Second	Third	Fail	Total
2013	7	11	2	1	1	22
2012	5	26	2	0	0	33
2011	7	18	5	1	0	32
2010	5	13	3	0	0	21

Table 2: Distribution of HEM Marks, 2010–13

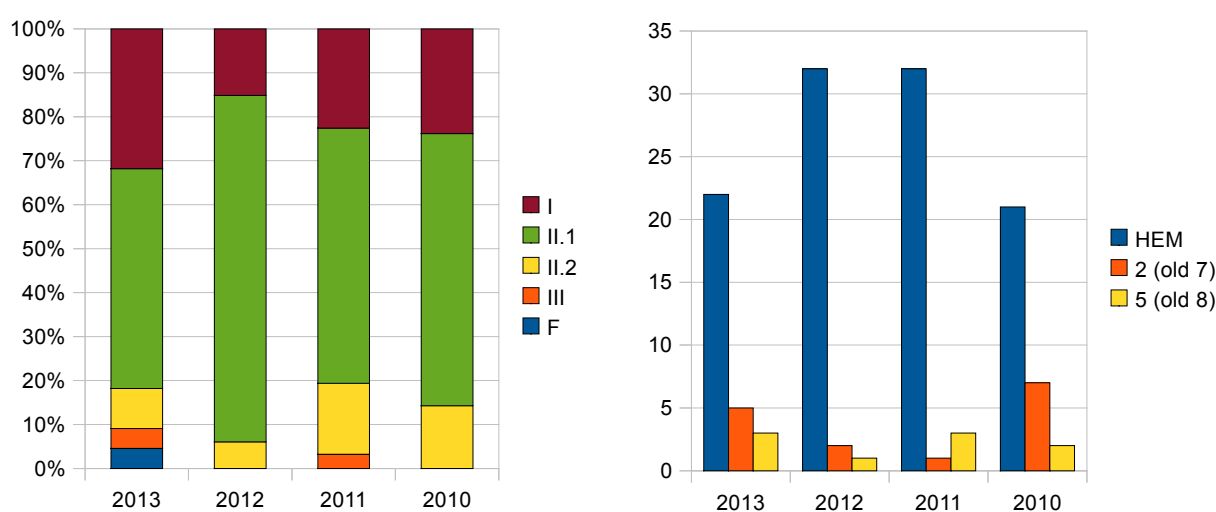


Chart 3 (left): Percentages of HEM classes, 2010–13

Chart 4 (right): Numbers of BBS candidates taking HPS subjects, 2010–13

### Class and mark distributions

The class and mark distributions for each paper are given in Table 3 and Chart 5 below.

Comments on performances on individual papers are given in the following section.

The examiners observed that all the First-class candidates had written dissertations, as indeed had 17/20 top-ranked candidates, compared to only 10/20 of the bottom-ranked ones. In other words, as noted in previous years' reports, Option A candidates continue to out-perform Option B candidates. However, that appears to be because the stronger candidates choose to write dissertations, not that candidates are disadvantaged by choosing not to. To check this I calculated the average marks of each candidate, excluding the dissertation, and compared them to the overall marks actually awarded. There is 98% correlation between the rankings with dissertation marks included and those without. In other words, we should not necessarily be encouraging every student to write a dissertation (although I note that our External thinks otherwise).

<i>Paper</i>	<i>First</i>	<i>Upper Second</i>	<i>Lower Second</i>	<i>Third</i>	<i>Fail</i>	<i>Total</i>	<i>Mean</i>	<i>Median</i>	<i>SD</i>	<i>Max</i>	<i>Min</i>
PS Essay	12	24	2	1	1	40	65.2	65.5	7.7	79	34
Dissertation	12	12	3	—	—	27	67.9	69	6.1	83	56
Paper 2	—	12	4	—	—	16	63.2	63.5	3.6	69	57
Paper 3	1	5	—	—	—	6	67.7	67	2.6	72	65
Paper 4	5	9	2	—	—	16	66	66	5.5	73	54
Paper 5	2	13	9	—	—	24	62.3	62	5.5	74	53
Paper 6	4	7	2	—	—	13	66.8	65	5.4	75	58
Paper 7	7	14	6	1	—	28	64.1	64.5	7.5	77	47
Paper 9	3	3	2	1	—	9	64	66	9.3	75	46
Paper 10	3	7	—	—	—	10	67.3	67	2.6	71	63
Paper 11	2	8	1	—	—	11	66.2	67	4	72	59

Table 3: Class and mark distributions by paper for HPS Part II: numbers of candidates

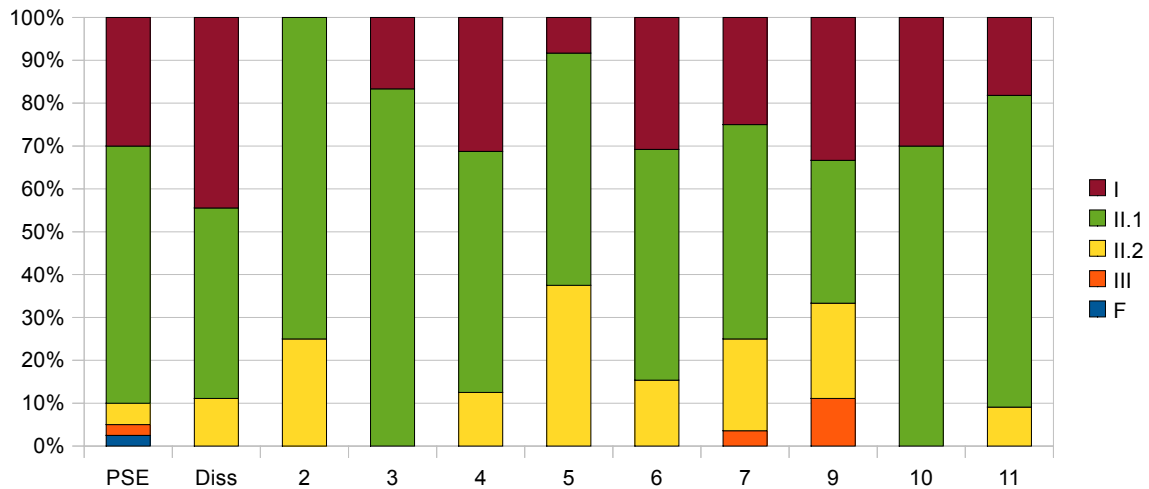


Chart 5: Class distributions by paper for HPS Part II: percentages of candidates

This year 21 of the candidates were female, 19 male. Men significantly outperformed women, with all but one Firsts going to men and all but one Thirds going to women. Men’s marks averaged 66.4, where women’s marks averaged 64.2. Only in the Primary Source essays did women outperform men, averaging marks of 66.6 compared to men’s average marks of 63.8. This was a similar profile to 2011 (when I was last Senior Examiner) — see Table 4 — but in 2012 women were reported to have outperformed men.

	<i>Firsts</i>	<i>Upper Seconds</i>	<i>Lower Seconds</i>	<i>Total</i>	<i>Mean Mark</i>	<i>Men’s Mean Mark</i>
2013	1/ 7	17/29	3/4	21/40	64	66
2012	8/16	—/20	—/1	17/37	68	67
2011	4/11	12/25	2/3	18/39	66	68

Table 4: Women’s class marks at HPS Part II, shown as a fraction of the whole, 2011–13

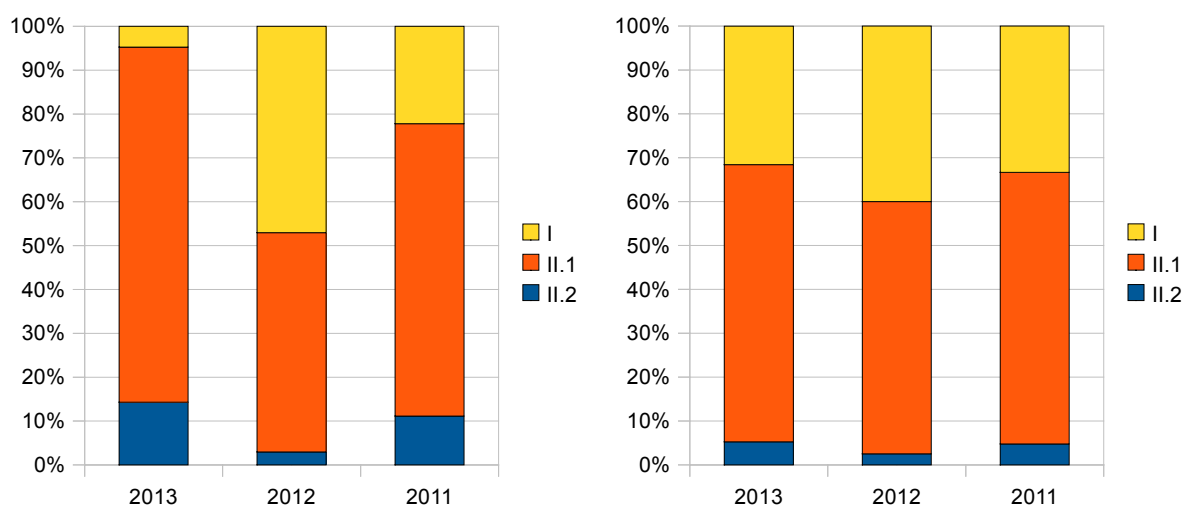


Chart 6 (left): Percentages of HPS Part II classes for women, 2011–13

Chart 7 (right): Percentages of HPS Part II classes for men, 2011–13

As can be seen in Charts 6–7, men’s performance has been much more stable over recent years (compare too with Chart 1 above).<sup>1</sup> Some further investigation should be carried out into whether this is simply a matter of small sample fluctuation or represents an underlying trend.

### Examining practice

Examination questions were set at the examiners’ meeting in Lent Term, following consultation with supervisors, lecturers and paper managers. The External Examiner—Staffan Müller-Wille, serving in his second year—also provided valuable comments on the draft papers.

As always, all elements of the course—dissertations, Primary Source essays, exam papers—were blind double-marked. Examiners were only rarely unable to resolve differences in their independent marks, so just one exam script and one Primary Source essay were given to the External Examiner for adjudication. In addition, he was asked to review and re-mark one exam script taken in unusual circumstances. The External Examiner was also given coursework and scripts from the upper and lower ends of the mark range, and those of the candidates on either side of the II.1/I boundary, for calibration. Finally, he was also asked to comment on the general consistency of marking by the examining team, and their relation to national standards.

The very late publication of the exam timetable meant that it was impossible to submit question papers on schedule for reproduction, although they were prepared in advance as much as possible, so that they could be handed in within a day or two of the dates being announced. There were no problems reported with the conduct of exams this year.

This was the first year in which subject-specialist Assessors were included in the examining team. Seven subject-specialist Assessors were recruited, six from within the postdoc community in the department and one from the Division of Archaeology (to co-mark Paper 11). They did not attend the Lent Term question-setting meeting or the Final Examiners’ Meeting.

<sup>1</sup> Not knowing the sex of the single candidate who obtained a II.2 in 2012, I have counted this as 0.5 candidate for both men and women.

The advantages were significant. The core team of Examiners could be reduced in number from six to five, while also reducing the amount of marking each did (on average 20 Primary Source essays, 7 dissertations, 45 exam scripts). The Examiners found it helpful and stimulating to co-mark with subject-specialists, while the postdoc Assessors benefited from seeing how their (and others') lecturing and supervising shaped the candidates' performances. Their workload was roughly half that of the Examiners (on average 11 Primary Source essays, 3 dissertations, 21 exam scripts). However, there was an increased managerial load on the Senior Examiner. Other commitments meant that one potential Assessor withdrew at the last minute, while others were available to examine only coursework or only exam scripts. All required training, and the marking schedule had to be arranged so that Assessors always co-examined with an Examiner and never with a fellow Assessor. Nevertheless, the overall benefits to the quality and good humour of the examining process significantly outweighed the organisational costs and—subject to the training and co-examining provisos mentioned above—the examining team strongly recommends that the use of Assessors becomes normal practice at Part II.

I am extremely grateful to all the Examiners and Assessors, the External Examiner, and the HPS office staff, for their efficiency, good humour and collegiality in their work this year.

### **Comments on performance**

#### *Dissertations*

These were generally excellent, attracting mean marks of 68 (median 69) compared to 65 for overall marks (median 66.5). The examiners would like students to be aware, however, that they are not seduced by pretty pictures that are merely decorative, but expect illustrations and diagrams to do narrative or explanatory work in the dissertation. Likewise, they urge the Part II Manager to introduce some training on citation, footnoting, and bibliographical practices in Research Methods seminars, perhaps through reference to a comprehensive online resource such as the *Chicago Manual of Style* citation quick guide ([http://www.chicagomanualofstyle.org/tools\\_citationguide.html](http://www.chicagomanualofstyle.org/tools_citationguide.html)). In particular students are encouraged not to rely on secondary accounts, synopses or selective quotation of primary sources but to read and cite from the primary sources themselves wherever possible. Close paraphrases of other writers' works must also be acknowledged and footnoted/referenced. Students should ensure it is clear whether the primary source has been consulted directly, or instead accessed from a citation in a secondary source.

#### *Primary Sources*

Primary Sources remain much more variable in quality (mean mark 65, median 65.5), to a significant extent dependent on the source itself and the way it was taught. Examiners noted, for instance, that many essays on the Defoe, Laudan, Stopes sources clustered clearly into a few restricted themes, while those on Longitude papers, Babylonian mathematics, and the IBBC report, for instance, tended to be much more varied and original in scope. In future years, the department is urged to choose Primary Sources, and develop seminars around them, that are

intellectually challenging and which encourage a diversity of approaches to the source.

### *Unseen papers*

More experienced examiners noted that answers to Section A questions have much improved in recent years. That is in large part down to the introduction of revision sessions that focus on techniques for dealing with them, and the department is encouraged to continue this practice. It may also be that the papers themselves are more coherent than in the past and thus lend themselves better to Section A questions.

However, the examiners noted that there was still more reproduction than generation in answers to both Section A and Section B questions. Much discussion was devoted to avoiding this problem in future years. Our recommendations focus on the exam-setting process in Lent Term:

- Paper Managers should avoid soliciting lecture-specific questions from the lecturers on the paper.
- Under no circumstances should supervision-specific questions be set.
- Paper Managers should be more proactive in reworking solicited questions, so that some invite candidates to draw on material from more than one lecture course in the paper, e.g., by comparing and contrasting materials, themes or approaches;
- Paper Managers should submit to the examiners both the draft paper and the questions as submitted by the lecturers, so that the examiners can check that appropriate modifications have been made.

The department may also wish to reconsider its policy about the restriction on the number of questions in section B and on the avoidance of disjunctive questions.

Poor handwriting continues to be a problem; we note that at university level the Teaching and Learning Services Support Group and the General Board Education Committee are currently exploring practical methods for implementing lap-top based exams.

### **Comments on specific papers**

These are best read in conjunction with the question papers. See also Chart 5 above.

#### *Paper 2*

There were no first-class scripts in Paper 2, primarily because candidates did not carefully consider the questions asked but tended rather to identify key words (e.g., 'authority', 'leprosy') and reproduce disquisitions on those rather than give nuanced, critical answers to what was being asked. In Section A, over half the candidates answered question 1, but took 'availability of sources' to mean only abundance, rather than considering different types of availability—physical, linguistic, electronic, etc. Nor did anyone acknowledge that new historiographies were not entirely source-led but could themselves lead to the identification of new types of sources or new ways of reading them. In Section B, the best answers to question 4 (answered by 10) distinguished historical circumstances and audiences for authority, as well writing versus practice. Likewise, in question 7 the best answers considered historians' perspectives as well as

the sources', and the physicality of symptoms versus the fear of them. The few answers to question 5 focused more on the importance for posterity than for late antiquity. Most responses to question 8 misinterpreted the question, either reading it backwards, or taking "Did ideas reflect treatment?" to mean "How did they do so?" There were very few answers to the Late Antique or Early Modern questions.

### *Paper 3*

Answers to question 1 made some attempt to pick out illustrative episodes, such as Lavoisier and Priestley, and some attempt to identify the emergence of singular and heroic authors, such as Newton. But there was no discussion of the kinds of institutions which produced conviction, such as academies, print or the reward system, and no discussion of the problem of assigning origins of ideas. Answers to question 3 were very varied, covering alchemy and the economics of chemistry, as well as the location and roles of naturalists, natural philosophers and their patrons.

Question 4 answers did not show a clear understanding of the distinction between nature and spirits implied in the question, and tended rather to produce brief histories of alchemy, with some odd errors. Answers to question 5 were in the main very strong, including interesting contrasts between natural history and natural philosophy, as well as some telling examples from taxonomy, experiment and chemistry. However, there was a wide tendency to understand the question as directed solely to diffusion. Question 8 produced some good answers, with details from Bougainville and Banks. In some cases they lacked a general argument about the wider implications of social evolution and stadial histories of social progress. Likewise, there were many good answers to question 9, which brought out Anglo-French contrasts between Kew and the Jardin du Roi, and made satisfactory references to collection as a commercial activity as well as an economically significant one, with some references to the high price of rarity in the specimen trade.

Other questions were answered only by one candidate (2, 6, 10, 11) or none (7).

### *Paper 4*

The 8 answers to question 1 were in the main very strong, drawing on many different themes to illustrate the means through which science changed, or was understood as changing society. The strongest picked out the distinction between reality and perception while the weakest used material from earlier centuries without offering detailed readings of Victorian material. The 5 answers to question 2 ranged widely, much weakened by a focus on the later 18th century. There were some good comparisons of the museum, the tropical island, the botanic garden and the teaching lab. But answers were much better at listing innovations than at explaining them, save through vague gestures at population increase and threat of popular uprisings. Question 3 prompted 3 rather good historiographic discussions and adequate reflection on the claim that disciplinarity justifies separate approaches, but some weakness around the question of disciplinary history as genre.

In response to question 4 there were 3 strong answers on the concept of professionalisation,

with interesting use of historiographic and modernist sources. There was a useful stress on the problems of anachronism and teleology, with apt examples from scientific societies and the notion of examining and training. Most of the 4 answers to question 5 focused exclusively on the story of cable telegraphy and electrical resistance. In the 6 answers to question 6, the ambiguity of the term 'progressive' caused major difficulties. It might mean that Darwin's theory described progressive development in nature; or that it had progressive social and political effects; or that it appealed to notions of progress. Answers were therefore widely distributed, mainly between analysis of Darwin's sources; or the use of the theory by racist, fascist and eugenic movements; or an account of the notion of degeneration.

The 13 answers to question 8 were very standardised: Lamarck, *Vestiges* and Darwin were all used to show the religious sources and analogies in evolutionary theory. Some more original answers noted explicitly secularist and atheist themes in vernacular evolutionism, but mainly retold the standard curriculum. Question 9 proved very difficult to specify. The 4 answers included references to the Royal Institution and the SDUK, some discussion of print media, and some discussion of exhibitions. But the main question—the relationship between the sciences and their media—was not centrally addressed. There was a lot of strong material amongst the 9 answers to question 10, albeit mainly standardised material from lectures. Not all candidates worked out that they were meant to address *either* geology *or* botany; almost all gave the same examples. There was, however, rare discussion of what 'empire' might have meant at this period. There was a lot of overlap with the 7 answers to question 11, as the matter of who took part in imperial science interacts with the question of how imperialism affected the sciences. There were good readings of much of the secondary literature, especially that on indigenous actors, and on the emergence of racism.

#### *Paper 5*

Answers to question 1 concentrated mainly on Chadwick and the Poor Law, though there was some discussion of pregnancy and rare but interesting references to Foucault and Illich. More on the historiography than the critique of the notion of medicalisation, which many candidates left undefined. Question 2 produced a generally weak set of answers, mainly a broad survey of germ theory and some passing references to the search for a magic bullet and to AISA. There was no real understanding of what the concept of risk was doing in the question. The quote in question 3 was mainly taken to invite comments on the effects of politics in medical decision, such as welfare measures and public health, rather than on the political implications of these measures. But stronger answers did survey a wide range of cases, including the implementation of the NHS and the debates about abortion.

More competent answers to question 3 understood religious and charitable institutions to have disappeared in the wake of the French Revolution, and associated the change to teaching hospitals and the clinical gaze with a theory of diagnosis and the disappearance of the 'sick man'. Some answers, however, did not answer the second question about the effect on medical theory. Question 5a provoked reasonable answers, looking at the relationships between antiseptics, germ



theory and surgical identity, though they did not centrally address the consequences of surgeons' hagiography of Lister. Most answers to question 6 recounted the straightforward tale of penicillin with most of the necessary details.

The answers to question 9a mainly took the line that women's campaigns for better perinatal health and for Sanger's model of birth control were as significant as the medicalisation of reproduction. There were good, solid accounts of the Puerto Rico trials of the contraceptive pill, and many answers that addressed feminist issues about abortion and about contraception. Responses to question 10 were rather broad, mainly concentrating on IVF and cardiology, with little discussion of media interest in scares or melodrama; there was much more on the way in which medics manage the news agenda and little on what news values might be.

#### *Paper 6*

This paper had remarkably successful answers and many candidates achieved first-class marks. In section A nobody answered the question on whether science is unified; there were 6 takers for question 1 and 7 for question 3. There was excellent performance on constructive empiricism questions (7 answers), mathematics as a language of science (4 answers) and biological kinds (5). The best answers on question 4 on observation-theory distinction (8 answers) also motivated the question explaining why it matters that there may not be a theory-free language.

#### *Paper 7*

The overall quality of answers was lower than in other papers. In particular there was quite a long tail of poor answers. Many candidates tended to air their own ethical views without supporting argument. Very few candidates answered either of the SSK questions; or the question on Bernal; or that on law. The questions set gave a very good coverage of all the material in the course.

In question 8 (19 answers) many candidates conflated the two important questions: is there a difference between genetic tests and medical tests? Is the difference ethically relevant? Many candidates recycled material they had already used when answering questions 2, 5 and 7. Question 5 and 7 prompted some poor answers which employed basically the same material as other candidates, often in confused form. Question 11 answers didn't discuss how different theories of well being might license different types and evaluations of human enhancement.

#### *Paper 9*

The answers were reasonably well distributed in Section A (3, 3, and 2); in Section B, apart from Q 11 which had no takers and Q4 which was the most popular, with 7, the distribution was quite homogeneous. The candidates showed a wide range of abilities. Most of them acquitted themselves well with the Section B questions (esp. on Locke and on Kant), but a number floundered on Section A. The best scripts were focused closely on the questions set and answered them with arguments, rather than reproducing materials from lectures, summarising articles, or reporting however large amounts of broadly relevant information.

### *Paper 10*

The performance on this paper was generally very high with many carefully constructed essays that queried the question asked and put forward an interesting thesis. The marks were between 63 and 71, with most candidates bunching in high 60s. Candidates did particularly well on the questions on Freud and Foucault and fairly good on questions on rational choice theory, laws in social science and philosophy of mind. Only one person answered a question on Keynes. On section A candidates flocked to the question on interpretation and the best answers tried to draw connections and contrasts between what Freud meant by interpretation and what interpretation might mean in anthropology and perhaps even philosophy of mind. As usual the best answers used lecture content as material for making up their own minds and showing independence of thought, rather than regurgitating.

### *Paper 11*

In Section A, most answers were given to questions 1 and 3. Responses to question 1 focused more on 'how' than 'why' historiography has changed, and many usefully used both celestial divination and mathematics as case studies. Answers to question 3 suffered from lack of consideration as to what was meant by 'cuneiform culture'. The 9 answers to question 4 largely reproduced lecture material, but the best distinguished between professional scribes and urbanites who needed a more limited range of education. Most responses to question 7 (7 answers) managed to distinguish causation and healing, but had little to say about different social contexts or historical change. There were many good answers amongst the 8 responses to question 8, which considered the reciprocal nature of the trust relationship and its pragmatic quality. There were some excellent answers to question 10, which reflected on both practitioners' and clients' needs and viewpoints.

### *History and Ethics of Medicine*

**History:** Question 1 was competently answered, with Galen offered as an important component but not well linked with Christianity. The focus was on plague and prayer with little attention paid to Early Modern changes. Most answers to question 3 considered both theory and practice, with heavy stress on Vesalius and Early Modern anatomy; and on French Revolutionary changes in doctrine. Question 4 attracted most answers in this section, mainly focussed on the French Revolution, the stethoscope and hospital medicine. Many students sought to incorporate their readings in Early Modern medicine too. Answers to question 5 made some links to Bentham, and there was much puzzlement about how successful the sanitary programme was, given that the question assumed its success. Question 6 was also popular. The focus was on TB and diabetes, with some attention to anaemia, mainly confined to stories of revised sense of these diseases, not on how much this has changed the *experience* of disease.

**Ethics:** This was a very popular section with most students choosing to answer the maximum three questions allowed. Almost everybody answered the question on abortion; organ sale was also a popular question. But the essays were generally of lower quality than expected. The main

problem is the perception that normative questions are very different from factual in that the answers to them are subjective, relative to taste and not amenable to genuinely sustained arguments. This is false: whether or not there is such a thing as moral truth does not take away from the necessity to genuinely consider the existing arguments on a controversy and to arbitrate between them responsibly and carefully. Notably students were better able to do that when they weren't emotionally invested into a debate. Thus abortion essays appealed to taste, emotion and subjectivity more than the essays on genetic testing. The key is not merely to go over the standard arguments and distinctions (for example, the violinist analogy), but rather to really show how the student was led to conclude whatever she concluded; rather than just using the existing arguments opportunistically to justify one's own views. Another common problem is reproducing lecture notes without explaining them properly. Thus students often drop names of philosophers without explaining their positions. Students should remember that their examiners are most likely not their lecturers, so the key to a good essay is not to show that they've been to lectures but to use lecture material to construct their own essay separate from whatever the lecturer did in the lectures.

### **Summary of recommendations**

#### *Assessors*

Subject-specialist as examiners should henceforth be used as a matter of course, subject to:

- a light-touch training session with the Senior Examiner for new assessors,
- Assessors always co-examining with an Examiner and never with a fellow Assessor

#### *Paper setting*

- Paper Managers should avoid soliciting lecture-specific questions from the lecturers on the paper.
- Under no circumstances should supervision-specific questions be set.
- Paper Managers should be more proactive in reworking solicited questions, so that some invite candidates to draw on material from more than one lecture course in the paper, e.g., by comparing and contrasting materials, themes or approaches;
- Paper Managers should submit to the examiners both the draft paper and the questions as submitted by the lecturers, so that the examiners can check that appropriate modifications have been made.
- The department may also wish to reconsider its policy about the restriction on the number of questions in section B and on the avoidance of disjunctive questions.

#### *Research Methods seminars*

The Part II manager should introduce some training on citation, footnoting, and bibliographical practices in Research Methods seminars, perhaps through reference to a comprehensive online resource such as the *Chicago Manual of Style* citation quick guide. In particular students are encouraged not to rely on secondary accounts, synopses or selective quotation of primary sources

but to read and cite from the primary sources themselves wherever possible. Close paraphrases of other writers' works must also be acknowledged and footnoted/referenced. Students should ensure it is clear whether the primary source has been consulted directly, or instead accessed from a citation in a secondary source.

*Gender difference in exam attainment*

Some further investigation should be carried out into whether the relatively low performance of female candidates in some recent years is simply a matter of small sample fluctuation or represents an underlying trend.

ER

16 June 2013