

NST Part IB History and Philosophy of Science

Senior Examiner's Report

8 August 2011

1. Summary of the Examination Process

As in previous years, the Part IB HPS examination consisted of two papers: History of Science, and Philosophy of Science. The examiners were Dr Salim Al-Gailani, Dr Kevin Brosnan, Prof. Hasok Chang (senior examiner), Dr Sacha Golob, Dr Jennifer Rampling, and Prof. Simon Schaffer. There was no external examiner this year. Special thanks are due to Dr Al-Gailani, who agreed to step in at the last minute as Dr Nicky Reeves had to withdraw due to medical circumstances.

Drs Al-Gailani, Rampling and Schaffer read the History of Science scripts, and Drs Brosnan, Chang, and Golob read the Philosophy of Science scripts. Each script was blind double-marked; of each paper, each examiner read 2/3 of the set of scripts, the rota being arranged so that there was an equal amount of overlap for each pair of examiners. A numerical mark out of 100 was given by each examiner to each script as a whole, with no breakdown of marks to individual questions.

The History of Science exam took place on Monday 6 July 2011, and Philosophy of Science on Tuesday 7 July. All HPS examiners met on Tuesday 14 July to agree all marks and discuss any issues. The examiners are to be commended for working against this tight timetable. The senior examiner thanks Dr Brosnan for attending in his place the NST IB meeting of senior examiners on 27 January, and Prof. Schaffer for similarly attending the final meeting of senior examiners on 22 June.

Most examiners reported difficulties in deciphering the poor handwriting of a number of candidates, for which inordinate amounts of examiners' time and effort were spent. This problem is likely to recur.

2. Candidates and Classes

There were 81 candidates, including 4 MIT students, 1 other junior-year-abroad student, 2 candidates for Part II Education Tripos, and 2 candidates for Part II Physical Sciences. The statistics given below are only of the 79 candidates who completed both examinations (1 candidate withdrew from both papers, and 1 candidate withdrew from the Philosophy paper).

The overall marks initially agreed by the HPS examiners (the 'raw marks') had the following distribution:

3 candidates	70.0% or above (1 st)
51	60.0% to 69.9% (2.1)
24	50.0% to 59.9% (2.2)
1	below 50.0% (3 rd)

This distribution, with 68.4% of the candidates in the 1st and 2.1 ranges, failed to meet the NST Part IB requirement that 60% and only 60% of candidates in each

subject area receive a mark of 60.0 or above (with a 2 percentage-point margin of deviation allowed in cohorts of 40 or larger). HPS also did not qualify for the exemption available to cohorts with high overall Part IA performance (the 'C value', which was 36.6 for our cohort). Therefore the raw marks had to be lowered according to the NST formula for piecewise linear scaling.

The distribution of the scaled marks, submitted to the NST Part IB final senior examiners' meeting, is shown in the table below alongside the results from previous years. Scaling had the effect of lowering all candidates' marks. The maximum amount of lowering was 2 points, and occurred at the 1st/2.1 boundary, bringing the raw mark of 62% down to 60%.

Class	2011	2010	2009	2008	2007	2006	2005	2004
1 (70-100)	1 1%	9 13%	15 18%	7 12%	16 27%	13 19%	13 13%	9 9%
2i (60-69)	48 61%	34 48%	37 45%	27 46%	22 37%	29 42%	46 47%	56 52%
2ii (50-59)	29 37%	24 34%	26 32%	25 42%	22 37%	27 39%	35 35%	40 37%
3 (40-49)	1 1%	4 6%	3 4%	0 0%	0 0%	0 0%	4 4%	0 0%
Fail (0-39)	0 0.0%	0 0%	1 1%	0 0%	0 0%	0 0%	1 1%	0 0%
Total	79	71	81	59	60	69	99	107

Since 2005, senior examiners' reports have included a gender breakdown. The breakdown for this year's examination is supplied with the previous three years for comparison (two candidates did not have their gender specified in the information we received, so they are omitted from the count):

Class	2011		2010		2009		2008	
	Male (% of total)	Female (% of total)	M	F	M	F	M	F
1	1 1%	0 0%	5 7%	4 6%	10 12%	5 6%	2 3%	5 9%
2i	26 33%	21 27%	18 25%	15 21%	22 27%	15 19%	11 19%	16 27%
2ii	14 18%	14 18%	15 21%	9 13%	12 15%	13 16%	11 19%	13 22%
3	1 1%	0 0%	0 0%	4 6%	2 2%	1 1%	0 0%	0 0%
Fail	0 0%	0 0%	0 0%	0 0%	0 0%	1 1%	0 0%	0 0%
Total	42 53%	35 44%	38 54%	32 45%	46 57%	35 43%	25 42%	34 58%

This year's distribution shows no significant imbalance of performance between female and male candidates overall.

3. General Comments on Performance

There are some notable features of this year's results.

(1) Comparison of performance in the two papers

Overall, the performance on the Philosophy paper was better than on the History paper. In raw marks: the average mark was 63.7% in Philosophy, and 61.0% in History; 16 candidates received first-class marks in Philosophy, and 6 in History. We have no clear explanation for these differences.

(2) Uneven uptake on questions

There was a great imbalance in the number of candidates attempting various questions, particularly in Philosophy (see exact tally in Section 4). In Section B of the Philosophy paper 5 of the questions received fewer than 10 answers each, indicating that just half of the available questions were really used. In Section A of the History paper, Q2 might have seemed to many students ambiguous and not so clearly connected to the course material. Some interesting, though unsurprising, tendencies were observed:

- In Philosophy, most candidates avoided questions based on a new course introduced this year (Philosophy of the Physical Sciences). However, many of them used materials very relevant to these questions in Section A (e.g., Philosophy Q6, Q8, Q10, in relation to Q1), indicating that they did absorb something of the course content but did not feel confident to answer entire questions on its basis.
- Candidates tended to go for questions that are clearly anchored to recognisable names, such as Darwin (History Q8), Popper and Kuhn (Philosophy Q5).
- Questions that demanded links between different topics were generally avoided.
- On the Philosophy paper, questions on general philosophy of science were strongly preferred to questions on the philosophy of particular sciences.
- On the History paper, both 20th-century (Q9, Q10) and ancient (Q12) topics were largely avoided.

(3) Overlap between questions

There remain some issues about the extent to which similar or identical material can be used in different answers. Future examiners need to concentrate, once again, on ensuring the widest possible spread of topics in questions. There was a particular difficulty with Q7 (second option) and Q11 on the History paper.

(4) Small number of first-class marks given

Only one candidate received an overall first, while the proportion of candidates receiving 2.1. marks was the highest-ever since 2004 (and the latter figure would have been even higher, without scaling). The paucity of firsts is a clear concern for students as well as those who have taught them. There is also a concern that this may result in high-performing students avoiding HPS at IB.

In the senior examiner's view, the low number of firsts this year is not due to any fault that can be assigned to anyone: lecturers, supervisors, examiners, or the students themselves. What is the explanation, then?

- The NST-dictated scaling of marks reduced the number of firsts, from 3 to 1; however, it may also be pointed out that 3 out of 79 is already a low proportion, and lower than in other years.
- A very large number of candidates chose questions that were perceived to be safe, and in answering those questions stuck too closely to the lecture notes and were not willing to present their own thoughts; this tended to produce 2.1 marks rather than firsts.
- Another pronounced tendency was that many candidates performed unevenly between the two papers. Even in raw marks (before scaling), only two candidates got firsts on both papers. In contrast, as many as 18 candidates scored firsts in one paper but not in the other; 6 of those even received 2.2 marks in the poorer paper. This unevenness contributed to the large number of candidates with overall marks in the 2.1 range, as the overall HPS mark is computed as a simple average of the marks from the two papers.

Therefore there are perfectly understandable reasons for the low number of firsts; what is still not explained is why the tendency for marks to lump in the 2.1 range was accentuated this year so much more than in previous years. It may be useful to seek information from previous examiners in this regard.

Setting explanations aside, we must explore ways of ameliorating this problem. Specific and urgent recommendations from the senior examiner are set in boldface below:

- If the NST scaling formula remains in force, it will always have the effect of depressing all marks whenever our raw marks put more than 60% of our candidates in the 1st and 2.1 ranges. As it seems likely that we will routinely have a high number of 2.1 performances, in order to prevent firsts getting scaled down to high 2.1's, we really must seek to overcome the humanities custom of not easily awarding marks much higher than 70%. If it is impossible to re-train the instincts of the examiners, **some agreed procedure for guiding the initial assignments of raw marks may be required, such as agreeing to award 80% (or higher) to the top performance each year and assigning the rest of the first-class marks accordingly.**
- The problem may be partly peculiar to IB, especially given the absence of the same issue at HPS Part II. One factor is that students get a choice of papers at Part II, but not at IB.
- Our standards of what constitutes a first at IB may be unreasonably high, given the nature of Natural Sciences students and given the fact that this is an introduction to our subject taken while the students are also coping with two other subjects. The question may be put as follows: is a flawless yet unoriginal reproduction of course content good enough to earn a good first? It would be, in many natural sciences subjects at IB level.

- **We should attempt to set the kind of questions on which IB students have a more reasonable chance of excelling, given their typical preparation and habits. It will be crucial for examiners to pay more heed to this matter when setting the questions.** This year's experience shows the wisdom of setting questions that are *both* obviously linked with straightforward course materials *and* with the kind of specificity as to force candidates to think freshly for themselves rather than simply reproduce standard materials (e.g., Q1 and also Q5 in History, and Q1 and Q5 in Philosophy). Such questions will be taken up by a reasonable number of candidates, and at the same time conducive to high performance by the better candidates. Meanwhile we must also emphasise to students that it is originality, not regurgitation, that earns a first; Natural Sciences undergraduates will not naturally see that point. These are matters on which lecturers and examiners should work more closely with supervisors.
- More fundamentally, the senior examiner's view is that we need to re-consider the contents and delivery methods of HPS Part IB. We must teach and train the students so that a good number of them *can* produce first-class work. An in-depth look at the content of HPS IB is on the horizon. One thing we must attempt is to prevent good IB students excelling in either history or philosophy while neglecting or giving up on the other. The History-side examiners this year felt that candidates were generally weak on historiography (e.g., not distinguishing between 'history of science' as 'what happened' and 'history of science' as a body of texts about 'what happened', even when explicitly invited to do so). This may be an important way in which IB students have generally not been able to make methodological lessons their own even if they are expressed clearly in lecture.
- Meanwhile, there are things we can do even as we stay broadly on the current syllabus. **It would be advisable to provide more specific focal points in lectures and supervisions, through which students can follow the larger issues; such focal points could be individual people, particular arguments, or salient examples.**

4. Comments on Individual Questions (in parentheses, the number of candidates attempting each question)

Paper 1: History of Science

Section A

Q1. If you were writing a book entitled the *Origins of Modern Science*, which historical period would it be about? (57)

The question produced some impressively critical reflections on the possibility of locating the 'origins of modern science' in any one period. Poorer answers were structured around a list of "greatest hits" of scientific advances without discussing historiographical issues, or providing any nuanced and fine-grained periodisation. Some strong answers engaged more specifically with the historiography of scientific revolutions. Science outside Europe was not discussed, apart from some references to ancient Iraq.

Q2. In what ways do earlier centuries' conceptions of the scientific role continue to inform our understanding of what it means to be a scientist in the twenty-first century? (23)

The relatively few answers to this question tended not to be strong. The question lent itself to rather vague and uncritical discussions of such allegedly 'scientific' characteristics as 'accurate', 'empirical' and 'objective'. Typically, answers began by identifying the traits of modern scientists and then looking for historical precursors. Few students addressed how historical developments actually shaped modern notions.

Section B

Q3. Compare and contrast the ways in which knowledge of the natural world was made by physicians and natural philosophers in early modern Europe. (23)

Few candidates took up the challenge of the question to compare (let alone contrast) natural philosophers with physicians, and none recognised that physicians might also practice natural philosophy. Only a small number of answers mentioned instruments. Most focused on anatomists rather than physicians, and few explored sufficiently the contexts of early modern medicine. Many answers to this question were descriptive rather than analytical.

Q4. Did experimental philosophy displace the magical and occult arts in the seventeenth century? (38)

Answers usually recognised that the relationship between experimental and occult philosophy was not one of simple 'displacement'. But few defined or specified the magic and occult arts, and 'magical' and 'occult' were usually lumped together. There was also little recognition that various practices (alchemy, astrology, magic) might be carried out by the same people. The best answers recognised that occult properties were a legitimate subject of experimental philosophy.

Q5. What was the role of novel instrumentation in the development of electricity and of chemistry during the eighteenth century? (19)

The relatively few candidates who answered this question generally did well, suggesting careful preparation of this topic. Good answers avoided the danger of equating novel instrumentation with scientific advances; the best extended their discussion to include novel adaptations of earlier instruments, and the capacity of instruments to extend what was previously observable. There tended to be too much focus on electricity, and on public spectacle.

Q6. Discuss this extract from a 1971 advertisement for the *Financial Times*: 'Isaac Newton is the British physicist linked forever in the schoolboy mind with an apple that fell and bore fruit throughout physics.' (22)

Students seemed confused by the question; there might have been better ways of posing a question about Newton's repute. Many candidates gave unimaginative biographies of Newton, or went into long digressions about the *Financial Times*, gender and science education in the 1970s, and public understanding of science in general. Better answers explored the history of 18C Newtonianism, using the

question as a starting point for investigating the popularisation and canonisation of Newton.

Q7. What differences did laboratories make to the development of the sciences during the nineteenth century? Discuss in relation EITHER to the physical sciences OR to the life sciences. (24)

There were some very well informed and detailed answers to this question. There was a preference for the 'life sciences' option, and many of those focused too narrowly on medicine, taking this as an opportunity to discuss the shift from bedside medical care to hospitals and laboratories (this also resulted in some digressions into the history of hospitals). Many students answered both B7 and B11, some dividing material between two and others making repetitions; this posed a difficulty for examiners.

Q8. Was Charles Darwin a professional scientist? (48)

Perhaps unsurprisingly, this was the most popular question. However, there were many weak answers, which described the career of Charles Darwin without answering the specific question posed. These tended to give unimaginative biographical accounts of Darwin and his career (the examiners suspected regurgitation of generic 'Darwin' revision material). Poorer answers did not interrogate the term 'professional', and often took it to mean 'expert'; few grasped that this was a question about the history of the making of professional science in the late nineteenth century.

Q9. Did Alfred Wegener's theory of continental drift initiate a scientific revolution? (5)

The low uptake on this question was disappointing especially considering that this was the one topic on which coordination between the two papers was attempted. There was a wide variation in the quality of answers. The best engaged with the delayed acceptance of the theory, but there was generally an inadequate discussion of the history of geology before Wegener.

Q10. What effect did the discovery of the structure of DNA have on the biological sciences? (13)

Few candidates approached the question of 'discovery' critically. Better answers challenged the assumption that DNA had an immediate or straightforwardly measurable effect on the biological sciences. Several discussed events leading up to the discovery of DNA, rather than consequences as required by the question.

Q11. 'It is of no significance whatsoever *where* nineteenth-century medical science was done. What matters is *how*.' Discuss. (34)

This was a popular question, and answers tended to be similar to each other, usually invoking French hospitals and labs, Koch, and Pasteur. Most candidates grasped that this was a question about hospital and laboratory medicine, and the best answers also addressed the geographical location of medical science. Weaker answers failed to link 'where' to 'how'. Generally there was surprisingly little reference to commercial routes and the availability of materials. (See comment on Q7, regarding overlap.)

Q12. Compare and contrast ideas about the origins and purpose of science in ancient and medieval Iraq. (14)

There were some good answers to this question, suggesting that the topic had engaged those students' attention. However, few answers displayed sufficient command of the historiography by, as invited, commenting on 'ideas about the origins and purpose of science'. There were some good discussions on the context of mathematics in ancient and medieval Iraq.

Paper 2: Philosophy of Science

Section A

Q1. When theory and observation clash, is it always theory that is to blame? (49)

This was a popular question, and elicited some high-quality answers; a high proportion of the answers were first-class performances. The question allowed for a wide range of material to draw from, and also provided a manageable challenge of synthesis; on both counts many candidates responded very well. Some answers displayed a satisfying degree of originality.

Q2. How, if at all, does scientific reasoning differ from everyday reasoning? (30)

Overall, performance on this question was visibly less good than on Question 1. Many of the interesting answers to this question were quite disconnected from the course material, so they were difficult to evaluate. It may have been useful for the question to include pointers indicating which topics from the syllabus would have been useful to address, as many candidates seemed quite unsure about this.

Section B

Q3. What is the best reason for thinking that the future will resemble the past? (51)

This was the second most popular question, with a good proportion of first-class answers. Yet very few answers were inspiring. Many were highly formulaic, exactly following the flow of ideas in Tim Lewens's lecture notes on the justification of induction. Credit is due to Dr Lewens for having imparted this material so effectively, but the question needed a bit more of an edge to shake students off from the well-trodden path. There were also too many people stating that reliabilism or pragmatism was the answer, without showing much convincing thought behind that verdict. Not many displayed an in-depth understanding of the reliabilist line.

Q4a. Do laws of nature make things happen? (14)

Q4b. Do explanations cite causal information? (28)

These questions were handled in a similar way as Question 3, except that the general quality was somewhat lower. Some candidates seemed confused by the way Question 4a was phrased. On Question 4b, some got distracted into discussing various difficulties of the deductive–nomological (D–N) model, without using those difficulties as a way to motivate the causal model.

Q5. What is the most significant point of contrast between Popper's and

Kuhn's views on scientific method? (59)

This was the most popular question, and differentiated stronger and weaker candidates effectively. Even weaker students must have felt able to take up this question, as it was easy enough to discuss some basic points about Popperian and Kuhnian philosophies of science. A number of candidates gave high-quality answers, with good accuracy and detail. Asking about “the most significant point of contrast” naturally encouraged students to exercise their own independent judgement, which some of them did very successfully.

Q6. Discuss cases of productive counter-induction. Are they rare exceptions or rather typical in scientific practice? (9)

Disappointingly few candidates took up this question (though many more discussed counter-induction as part of their answers to Question 1). The quality was mixed, between strong candidates who felt confident enough to tackle this question and those who must have wandered into it in the absence of anything better they could offer on other questions.

Q7. Could a false theory or model be successful? (27)

Most candidates took this as a thinly-veiled question about the realism debate, which was the correct reading of one aspect of the examiners' intention. However, few got into any in-depth discussion about the nature of models (or theories), which we had also hoped for. Also, at the surface level the answer to the question may have been an all-too-obvious 'yes'; it might have been useful to phrase the question in a more nuanced way.

Q8. Have there been any major theory-changes in the physical sciences which exhibit Kuhnian incommensurability? (6)

We would have expected a question on Kuhn to be popular, but the demand to produce examples from the physical sciences may have been too formidable for most students. This was also an example of a question that clearly reached across two courses (Methodology, and The Philosophy of Physical Sciences), for which students would not have been able to prepare through supervision essays.

Q9. Should metaphysics change according to the latest development in science? Discuss with reference to examples from the physical sciences. (8)

This question was probably too difficult, as most candidates are not well-equipped in metaphysics or in the physical sciences. But some answers were excellent.

Q10. Is measurement theory-laden? How does that issue relate to the Duhem-Quine thesis? (4)

This may be another case of avoidance of a question reaching across two courses. Many candidates did discuss theory-ladenness and the Duhem-Quine thesis at a good level in their answers to Q1, but it may have been too difficult to connect those themes with measurement. It is not clear how well the philosophical issues about measurement were conveyed to the students in lectures (or to supervisors).

Q11. 'Evolutionary theory is not scientific because it is no more than a

tautology.' Discuss. (7)

The low uptake on this question was disappointing, and less understandable than the other cases. The quality of answers was also generally not very high. It is a matter for concern if students in Philosophy of Biology failed to grasp the basic worry that the “survival of the fittest” may be a tautology.

Q12. Must scientists always tell the truth? (24)

Like Q2, this question was difficult to mark as candidates could return relatively cogent and plausible answers without engaging much with the course material. Some of the candidates, while making various good points, failed to synthesise them into a cogent whole.