

Dr. Staffan Müller-Wille
College of Humanities
University of Exeter
Amory Building, Rennes Drive
Exeter, Devon EX4 4RJ

Exeter, 17/06/2012

To the Vice Chancellor
University of Cambridge

External Examiner Report, NST Part II History and Philosophy of Science (HPS) and NST Part II (BBS) History and Ethics of Medicine (HEM), 2011-2012

This was my first year as External Examiner for the two courses mentioned above. Examinations consisted in three (four) unseen examination papers, two primary source essays, and a dissertation (optional) for HPS, and one examination paper and a dissertation (optional) in HEM. Essays and dissertations were made available to me in good time and electronically, which allowed for a quick, but thorough turnover. Internal assessment was carried out by two independent examiners for each element. I was asked to resolve non-agreed marks in seven cases, and to calibrate marks for a set of papers representative of classes, at the borderlines of classes, or resulting in significantly different internal marks. I did this in all cases by marking the papers myself. I was also given the opportunity on the day before the Examiner's Meeting on June 12 to have a look at all the examination papers to make vertical and horizontal comparisons of marks and to scrutinize cases in which results of an individual student were uneven.

In the following, I will first make some general comments on student achievement and the examination and assessment procedure as a whole, and then add some comments and observations on individual elements of the examination. I do not see any major problems with teaching and assessment in both courses, and my remarks below are solely made for the purpose of enhancing future teaching, assessment, and student experience.

Student achievement in NST Part II HPS was impressive, which is not surprising, as the Cambridge HPS programme certainly attracts the best and most enthusiastic students in the field. The level of solid and detailed knowledge and understanding acquired in the course is high, and coursework demonstrates that students also gain significant skills in independent research during the course. This should be seen against the background of a programme that is already quite challenging in content, dealing with complex subjects and placing a lot of emphasis on critical reflection. In general, I would judge that achievement is well above the average in comparable programmes I am familiar with (Exeter and Leeds). In HEM, students achievement fell within a much narrower band of high, but usually not excellent performance, which is in line with previous years. Again, this is to be expected, as this course is taken by students in the biomedical sciences, which tend to reproduce taught content, and tend to have less developed skills in the critical analysis of texts that would allow them to produce coursework of high originality.

Statistically, there is a clear tendency to avoid high Firsts and the lower end of the II.1 and the II.2 range of marks, a problem that seems to continue from previous years. More than half of the marks of primary source essays, e.g., are low Firsts and high II.1s (60% at 67+). In my own marking of roughly 30 papers, however, I did not reach significantly different conclusions, with a few notable exceptions of far above or far below average achievements. In these cases, it turned out that internal examiners had themselves made written comments on coursework that clearly justified a higher or lower mark respectively if compared with the existing marking criteria. My

recommendation is therefore that examiners force themselves to look at marking criteria again and again, and apply them “mechanically”, in a check-list manner, to course work. Another way to avoid inflation of Firsts in particular would be to introduce a new class of “starred Firsts” for essays scoring marks of 80 and above. With respect to II.2’s, I encourage examiners to recognize that work falling in this class is described as “competent and broadly relevant” by the marking scheme, so that there is no reason to reserve this class for clearly confused or deficient work only.

The assessment procedures in place, with two internal examiners marking papers in a “double-blind” manner, is labour intensive, but guarantees a maximum of fairness. Inevitably, such a procedure will generate strong discrepancies between markers in some cases. Most of these marks were agreed internally, sometimes by consulting a third internal examiner. The remaining cases of non-agreed marks were resolved by me, after careful scrutiny of the paper in question. No consistent pattern emerged that could have explained the sometimes pronounced discrepancies. In each case, it seemed that examiners had attended to one particular aspect of the marking criteria in isolation, again a problem that can be avoided by consulting the marking criteria in a more “mechanical” fashion. A point of good practice I would like to highlight were the often detailed and substantial comments that examiners provided on primary source essays and dissertations with their marks. These are currently not made available to students, and I would recommend doing that in the future. It adds to the transparency of the marking process, and provides useful feedback. Another, minor recommendation is to take into account formalities like referencing style, punctuation etc. more explicitly in marking, because attention to these matters is expected from professional environments.

I would like to add some further comments on individual elements of the examination. In marking primary source essays in HPS, I noticed that there was less evidence for students applying critical techniques in reading sources than one could perhaps wish for. Most essays just went ahead and interpreted sources for content, and they were judged by examiners on that account also. As the purpose of this element is acquisition of critical skills, I suggest that lecturing puts more emphasis on archival, philological, and statistical methodologies.

It was student achievements in dissertations that impressed me most of all. I think, however, that with 20% too little weight is given in the overall marking to what after all is the result of a sustained effort to produce a piece of independent research. There were a few cases, in which students had excelled in their dissertation, while performing not as well in the other elements. I would also suggest to the department to reconsider its recent decision to make dissertations an optional element.

With respect to exam papers, I found marking a bit uneven in the cases where I compared exam scripts “vertically” between students who had received the same marks. This is certainly due to the huge work load that this element creates for examiners in a relatively short span of time. One way to solve this would be to move to moderation in marking exam papers, i.e. having second markers look at a set of first marked papers in comparison. There were also quite significant “horizontal” differences between average marks achieved on different papers (especially papers 1 and 10). In hindsight I can say that this is probably due to the quite different styles in which questions were posed. More attention should therefore be paid in the future to homogeneity in the level and clarity of questions across papers.