

REF 2014

So what did we learn from the REF?

CUCD has been gathering feedback and opinions.

So have we all. These thoughts are not intended to duplicate the many local enquiries by senior management teams, or the flurry of responses that followed the publication of [the results](#), let alone some pretty hard hitting analyses in the Press. Instead these are simply a few observations about what REF did to our discipline, and what the published results mean for us.

Some of this comes out of discussions at Standing Committee and elsewhere, some is based on material in the public domain, and of course none at all derives from panel members, bound as they are to HEFCE by the most frightful *Omertà*. The formal feedback from the panel chairs has already been published, including [the Classics entry](#) (on pages 58-68) from sub-Panel 31, our panel. It is generally upbeat and helpful and gives a very positive account of the strength of research in classical subjects. It is clearly essential reading for those already preparing for the next REF about which we still know very little, not even the date, except for the announcement about [Open Access requirements](#). This article is less authoritative, a personal view from someone who has been close to the gossip and far from the work of REF 2014.

How we prepared

Preparing for REF felt like much more work than preparing for RAE.

One reason was the need to compile impact case studies, on which more below.

Another reason was the generous resources of cash and staff time which university senior managers were prepared to spend on mini-REFs, allegedly also on ghost-writers for environment and impact statements, and on external consultancy. Senior academics, former RAE panellists, (and in a few cases current REF panellists, although I did not hear of any cases in Classics) were recruited to vet drafts, to assign marks to outputs, and to advise on inclusions and exclusions. Anecdotal evidence suggests this was much more lucrative and much less work than conventional external examining.

The most expensive component of these dummy runs was the time and energy and nerves of the poor dummies being sent on their practice laps. Attempts are being made to estimate the total cost of REF2014: [one estimate](#) has it between £0.5 and £1 billion, perhaps around ten times the cost of RAE2008. Another estimate is that it cost around £1.2 billion: that is about the same amount of QR funding that is distributed each year by HEFCE. And then there is the opportunity cost: how much more of what was being measured could have been written if it was not being measured, or if it had been measured with a lighter touch?

Another cost is less easy to measure. Goodwill and collegiality. 401 individuals were submitted to REF 2014, that is 60 less than were submitted to RAE 2008. The figures look differently in term of FTEs with a smaller drop from 415 to 383. Yet CUCD statistics

show there was almost no change in the number of staff employed in Classics departments between 2008-9 and 2012-13 (the latest figures that are available). For what it is worth, the figures were 510 individuals in each year and a shift from 445 to 443 FTE. So there has been a significant increase in the number of classicists not submitted. This figure conceals big differences between departments. Some departments submitted everyone who was eligible, others deployed a deeper cut. At least some of the ranking must have been affected by strategies of exclusion (assuming the 'right' colleagues were excluded). Decisions on how inclusive to be will rarely have been taken at the level of departments: but judgments over precisely who to exclude much mostly have been made locally. CUCD should perhaps watch out for the long term consequences of these tactics.

What did we write?

Successive Classics panels have tried to define the subject as inclusively as possible. The sub-panel 31 report celebrates the diversity of subject matter and format in the submissions.

As before, the range of outputs was quite wide. Just over 30% were monographs and most universities asked for these to be double-weighted. Almost every request made (98.8%) was agreed to. Unless the rules change dramatically, that means we should probably all request that our monographs be double weighted next time.

Journal articles made up about 25% of submissions and book chapters just over 30%. The remaining 12% or so included edited books (8.4%) and various other kinds of outputs. Scholarly editions made up only 1.9% of the outputs submitted which might ring a few alarm bells.

How was it read?

We were lucky, once again, to have a sub-panel of our own.

The archaeologists had to cohabit with the geographers, and although our panel was the smallest it had an impressive range of expertise. It was good to see that most of CUCD's nominations for membership were taken up. Two recently retired colleagues were called on in the final stages to help meet the deadlines, and they generously agreed. As a discipline we should be grateful for the time panel members dedicated to reading and to producing carefully considered feedback. There is no need here to repeat the reports they made on the health of various subdisciplines. Those reports were broadly positive, not just about what has been achieved but about the future too, noting the emergence of new specialisms and the presence excellent submissions from early career researchers. The picture their report presents is of a diverse and vibrant research culture in UK Classics. All this is good news.

Everything submitted to sub-panel 31 passed through the hands of classicists. But not all classicists were submitted to sub-panel 31. A number will have been submitted to the History panel or perhaps Archaeology. This is why in the results for Panel 31 there are no entries for a number of CUCD members including Birkbeck, Cardiff, Leicester, Roehampton and Swansea. It is also why KCL submitted 29.9 FTE while UCL 13.00 FTE (UCL ancient historians being in history and their classical archaeologists in archaeology). As in previous exercises it was possible for panels to refer items they did not feel qualified to judge. It would be good to know how effective cross referral

really was. It is surprising to read in the report from Main Panel D that only 4% of outputs were cross referred. It is also clear some sub-Panels exported a lot more outputs than did others. On the face of it, it seems likely that many researchers from CUCD departments had some of their work assessed by other panels.

What else have we been up to?

The broad levels of grant income and its distribution look fairly similar to that of previous years. This perhaps unsurprising given the limited number of funders, and also how well we have done in the past. We still compare quite favourably with many other humanities disciplines. We are also producing more doctorates than ever. There are other questions that might be asked about this, and [some tough answers](#) have been proposed. To the credit of sub-panel 31 attention is drawn in their report to growing casualization, to the plight of early career researchers on temporary and/or teaching only contracts, and to the failure of some of us to think hard (or at least to write coherently) about how our research strategies take account of them.

What about Impact?

This was the major innovation of this exercise. Preparing for it devoured vast amounts of time and energy, partly gathering data we never knew we would need, and partly trying to understand the complex definitions and rules about eligibility. A few departments evidently fell foul of the latter, either because it was difficult to link the public engagement and outreach work they have been engaged in to specific research outputs, or because good examples were ruled out because staff had moved since the original work was done.

Life was toughest for small departments because each impact case study had such a disproportionate impact on the profile as a whole. There are stories of individual researchers being excluded not on the quality of their work, but because their inclusion would raise the size of a department to the point where another impact case study would be required. Much of this is invisible in the eventual results.

The final scores were on the whole less terrifying than many had feared. Overall 41.4% of impact case studies were 4*, compared to only 29.4 % of outputs. If 4* and 3* are added, as in some published tables, the contrast is even sharper: 88.2% of impact case studies were judged in the top two categories, as opposed to 70.4% of outputs. Perhaps there is not much point comparing such different kinds of assessment, except that they end up being worth the same when the final profiles are calculated arithmetically.

All 59 Classics impact case studies can now be read [on the HEFCE site](#) . They are quite varied, but nowhere near all of our subdisciplines are represented. A crude count suggests that archaeology featured as the main element in about half the case studies. A quarter made significant use of digital resources and about a fifth were based reception studies (the categories do overlap). Ancient philosophy featured in surprisingly few impact case studies. Finding ways to demonstrate the impact of philological research or literary criticism proved more difficult but a few case studies were based on Greek drama, and a few emphasised how research had had an impact on education in schools and universities. It only became clear part way through the period of preparation that producing pedagogical materials

or changing education in other ways was regarded as a legitimate form of impact (so long as it was not our own students who were benefiting from our research).

As the newest component of the REF, impact will probably change most in the next exercise. For now we know we need to gather data, and build impact into research projects from the start. Perhaps classics departments without archaeologists or digital humanists should hire a few?

Winners and Losers

RAE and REF have always been presented as based on absolute measures of research quality. There is no rationing of 4* grades. We could, in principle, all get top marks. The Classics panel did not feel that the 400 odd individuals whose work was assessed could easily be sorted into 4*, 3*, 2* 1* or unclassifiable researchers. This obviously has implications for the effectiveness of excluding individuals as opposed to choosing which outputs to submit. The profiles of every department had some 4* elements and most had a little 1* as well. Compared to many other subjects, Classics did not seem to have much of a tail.

All the same league tables were at once produced, based on 4*, 4*+3* or GPA and there has been a national debate over the whether research intensity, research power or market share is the better measure of virtue. One well informed former-VC told me that from looking at websites he reckoned there are now about 35 UK universities in the top 10. And depending on the ranking method adopted some universities shoot up and down the tables. University R is 38, 27 or 19 according to the measure applied, University B is 14, 35 or 34 by the same measures, and so on.

Classics rankings, for whatever they are worth, have been fairly stable through successive exercises. It is clear enough why this is. The cull of Classics departments during the 1980s and the effect of successive RAEs has removed low functioning departments altogether as well as encouraging universities to support their researchers better. Most classics departments are pretty similar especially in terms of the kinds of universities where they are located (meaning broadly similar workloads for staff and resourcing for research). In REF terms most of us inhabit similar environmental niches.

The only very obvious differentiation is that the largest departments rarely do very badly and the smallest rarely do very well, almost irrespective of the measure taken. No surprises here. Size brings larger library resources, more graduate students, better staff/student ratios and often a greater capacity to support research leave. Classics degrees are more complex to run than most humanities degrees (a wider range of subdisciplines, the need to teach languages, usually at a variety of levels) and so members of small departments often have higher administrative burdens than their colleagues in philosophy and history etc. We all know this already. The results of REF2014 had no real shockers for classicists. Only five departments submitted less than 10 FTE and they were near the bottom of most tables. The correlation was less clear at the other end, but no large departments did very badly. Middle sized departments were shuffled a little – between exercises and between tables – but only because these departments are so similar that any REF-based rankings are sensitive to the slightest variation.

Use and Abuse

More serious for Classics departments are two less legitimate means of comparison.

First, given the absence of much of a tail in our discipline – taught in around 30 universities, of which only 22 made submissions to Panel 31 – it arguably doesn't mean as much to be in the bottom quarter of the league table as it does for some other humanities subjects. The History panel looked at 83 departments, the English panel at 89. Being the median department in English is not the same as being the median department in Classics. All the same there are signs that, as in all previous exercises, the lower ranking Classics departments are being given a tough time.

Second, universities are already producing internal tables that compare the profiles of different departments within the same institution. One has to have a very high degree of faith in the comparability of the standards different panels applied to think this sort of comparison tells us very much. As a former RAE panellist I certainly do not have that faith in the absolute equivalence of gradings produced. Unfortunately many senior academics do seem to suspend their critical faculties when they join senior management teams, and collectively forget the limits of what the data-sets generated by the REF can tell us. We do not always help remind them of this. It is easy for us to give in to the temptation to make those rankings seem more legitimate and secure than they are when we boast of our temporary achievements in them. Nationally Impact Data and case studies are being deployed to show the cultural value of the humanities. Most of us believe in that cultural value (or just 'value') and most want to share what we discover and speak out for why it matters. But if we sign up to impact rankings as the best measure of value, we will have only ourselves to blame if less impactful activities are driven out of the Classics. The sub-panel itself was not convinced. To cite their report (p.63 ch. 28) "Sub-panel 31 remained doubtful whether what it was able to demonstrate in REF represented the true impact of different units adequately."

As always CUCD is glad to know of any threats to funding, departments or posts that might follow, and will keep that information confidential if necessary until the point when any public action seems helpful.

Greg Woolf

Greg.Woolf@sas.ac.uk