Bulletin of the

COUNCIL of UNIVERSITY

CLASSICAL DEPARTMENTS

Number 35
2006

CONTENTS

Chair's Report, 2005–6

Dr Johnson said, 'The only end of writing is to enable the readers better to enjoy life, or better to endure it'.[1] To what extent the Chair's annual report to Council meets either of those aims, others must judge.

Consultation, consultation, consultation

The high peaks of activity in a very busy year have generally been in the run-up to deadlines for the submission of consultation responses to agencies such as the DfES (Department for Education and Science), HEFCE (the Higher Education Funding Council for England) and the other organizers of the RAE, AHRC (the Arts and Humanities Research Council), and the QAA (Quality Assurance Agency for Higher Education).[2] The AHRC is the most prolific consulter. At the time of going to press, consultation deadlines were looming in connection with the second report of the Burgess Committee on degree classification, the AHRC's proposals on future funding of postgraduate study (CUCD has spotted an opportunity to press, with other subjects, for the possibility of two-year research training MAs), and the review of peer review by RCUK (Research Councils UK). Three other consultations are discussed in more detail below.

Frequent requests from Standing Committee for material to be incorporated into responses place a heavy burden on departmental Contacts and Heads. (We note with relief that the AHRC has simplified its nomination procedures for the Peer Review College and grant-giving panels.) It is evident that a marked degree of 'consultation fatigue' has set in, judging by the tailing off of the numbers of response received by the Chair. The situation is not helped by the way requests seem to cluster in vacations and at the start of the academic year. The Chair wishes to put on record, however, that the responses he does receive are usually detailed, invariably thoughtful, and always to the point.

The two consultations that generated the most heat were that by the AHRC (deadline in April) on the UK's prospective contributions to ERIH (the European Research Index in the Humanities), and that by the DfES (October) on metrics in the RAE. Both are discussed in detail elsewhere in this *Bulletin*, but it may be appropriate to put a few words of additional orientation on record.

European Research Index in the Humanities

ERIH is a project of the European Science Foundation ('science' is to be understood in the wider sense prevalent in Europe outside UK, embracing the humanities), the first phase of which was due to be completed this year. Its generally laudable aims include the production of a *research tool* that will display European research better than existing indexes do. But its apparent classification of journals in a qualitative ranking sparked fears (most strongly but not exclusively in Britain) that it could be used as a crude and misleading *research assessment* tool.

Early in 2006 the ERIH panel chairs, drawn from a range of EU states, had raised such concerns with ESF—and had been heard. It was unfortunate, therefore, that the subsequent Europe-wide consultation by ESF (issued via the AHRC, in our case) prompted by these concerns coincided with the unexpected declaration by the UK Treasury that metrics (including publication metrics) would wholly or largely drive future RAEs, and possibly the 2008 RAE. Understandably, many classical departments were not sufficiently reassured by the signs that ERIH would be modified to prevent them sending, through the Chair, the most negative response we know of by any UK subject. By a four-to-one majority, respondents wished CUCD not to participate in the compilation of a journals lists. That was also the position of most UK arts and humanities subjects, archaeology being one exception.

Although modifications to ERIH would have taken place without this mass protest, the strong reaction did force the organizers of future RAEs (chief among whom HEFCE) to rule out using data derived from ERIH in any future RAE metrics.

Department for Education and Skills

The spotlight then turned to the wider principle of research metrics. Just before we went to press,

---

[1] Corrected version from www.samueljohnson.com/apocryph.html, accessed 20 Oct. 2006.
[2] I give the expansions of these initials with an eye to readers in a few years' time, when no doubt they will have changed. For some reason, current usage varies in the use of the definite article in ordinary speech, as above.

the DfES gathered responses from UK stakeholders about the possible implementation of the Treasury's wishes. The signs are that responses to this consultation, too, will be overwhelmingly unenthusiastic. Readers will find excellent material in this *Bulletin* to focus future thinking. Whether the critical responses (particularly that from the British Academy) will be listened to is as yet unknown, though the HEFCE–AHRC panel on the matter (which includes two archaeologists), while not setting its face against any extension of metrics, does send a clear signal about the problems of using publication metrics in the humanities.

On a happier note, dialogue with the AHRC's highly qualified officers and its Chief Executive (Philip Esler, a New Testament scholar) has been active and amicable. The Chair takes every opportunity to remind colleagues that the AHRC, which spends *c.*£70 million a year on UK research in arts and humanities, is an ally, not the enemy. Much friction can be avoided, and misconceptions nipped in the bud, by picking up the phone to Bristol rather than composing splenetic e-mails.

Quality Assurance Agency

On another positive note, the Chair attended a QAA fact-finding meeting in December 2005, at which the Agency demonstrated that its approach is softer than it was a few years ago. The Agency (funded by the HE sector) does not regard its role as that of closely prescribing what we should and should not do, but that of facilitating the generation of good standards and the sharing of good practice. QAA staff present advised that if university administrators represent the Agency's guidance as mandatory, they may often be exaggerating in order to discipline departments. CUCD hopes this advice proves useful.

The benchmarking revision working party chaired by Professor Robin Osborne completed its work in spring 2006, and the draft of the lightly revised undergraduate benchmarking statement has been made available for comment by the QAA.

External relations

One incidental benefit of consultation overload, particularly during and after the ERIH furore, has been that arts and humanities subjects have talked to each other much more than before. At the prompting of the Chair in April, and again in June after the (now annual) AHRC Meeting with Subject Associations and Learned Societies, we set up an e-mail list of officers of UK arts and humanities bodies (chiefly with a Higher Education remit), which is now used regularly for exchange of views and sharing of draft responses. It is hosted by CUCD's website and will be updated regularly.

While different subjects do not, of course, have identical views on all matters, there is a remarkable degree of agreement. Frequent exchanges of ideas can only be a good thing for the defence of academic freedom and standards.

It almost goes without saying that we talk regularly to the officers and staff of the Joint Association of Classical Teachers, Classical Association, Society for the Promotion of Hellenic Studies, and Society for the Promotion of Roman Studies. We are developing links, informally at this stage, with associations abroad to whom we distribute printed copies of the *Bulletin* and with some of whom we correspond regularly by e-mail: for example, the American Philological Association, the American Institute of Archaeology, and the Greek and Latin Studies Committee of the Royal Irish Academy. CUCD members are encouraged to suggest other bodies with whom we should liaise, particularly those with a Higher Education remit or equivalents of the Classical Association.

All-Party Parliamentary Group for Classics

Another important UK link has come into being through the All-Party Parliamentary Group for Classics, set up at the suggestion of our own Peter Jones, your Chair, and Michael Fallon, MP. Michael, a Classics graduate of St Andrews, is chair of the group, which has a membership in both Houses and officers from all three main parties. It is early days, but we expect CUCD and other classical bodies to develop close links with APPGC, along the lines of the very active collaboration between the Society of Antiquaries of London and APPAG, the equivalent group for Archaeology.

APPGC, we hope, will provide not only a means of putting Classics in the public eye and countering threats, but also a forum for classical organizations to meet and discuss issues of mutual interest. It is not obvious that, say, four responses from 'Classics' on a matter concerning Higher Education (such as ERIH) are better than a single one from CUCD, the specific voice of

Classics in Higher Education. (It is worth noting that CUCD's list of nominations for the RAE panel was adopted pretty much unchanged by the funding bodies.) Nor should CUCD necessarily hold identical views about public policy to those of CA, JACT, or the two Societies. But two (or more) heads can be better than one; and though there are strong similarities and some duplication between, on the one hand, our Standing Committee, Contacts, and departmental heads and, on the other, the academic council members of the other bodies, it is entirely appropriate that on matters of interest to classical departments, for which CUCD has a distinctive remit, it should consult widely. We have, for example, been closely in touch with JACT in response to concerns about the OCR examinations board.

Other business

CUCD has liaised closely with the Institute of Classical Studies (whose director is an *ex officio* member of Standing Committee) regarding its current problems with the University of London. We stand ready to support the two Societies who are the principal stakeholders in the Library—though Standing Committee is mindful that it is the interests of classical departments, not the Societies, that we are mandated to uphold. As we went to press, negotiations appeared to be entering their most critical phase.

With the support of public figures including Tam Dalyell (until recently rector of Edinburgh University and Father of the House of Commons), we continued to press for action to resolve difficulties over the only classical postgraduate diploma in education in Scotland, hitherto taught at Jordanhill under the auspices of the University of Strathclyde and threatened as a result of funding decisions by the Scottish Executive. Happily, there are signs that something new may be put in place by Scottish universities.

Standing Committee prompted a discussion of important issues bearing on the management of early career staff, following an important paper by Dr Steven Green. While many of the issues are generic rather than subject-specific, there is no reason why, if there is a need for more sensitive handling of applicants and appointees by personnel departments, CUCD member departments should not take the lead

in encouraging good practice.

We organized a panel (co-sponsored by the ICS and the Classics Subject Centre) at the CA conference in Newcastle upon Tyne about publication issues of interest to research students, which was extremely well attended. We benefited from presentations by the classics editors of CUP, OUP, and Routledge as well as by Kate Pool of the Society of Authors, an expert on intellectual property rights. This session will be followed up at Birmingham in 2007 with a session on journals.

Past and future

It is a pleasure to close with renewed thanks to our supporters in public life, to all members of Standing Committee since 2003, but particularly to the one Treasurer (Patty Baker) and two secretaries (Philip Burton and Bruce Gibson) with whom I have worked. Their jobs have become increasingly demanding, and they have done them brilliantly. The same goes for those other colleagues, both elected and co-opted, who have undertaken special responsibilities under the Constitution.

My principal ambitions as Chair have been to encourage the widest possible involvement of departments in formulating responses to consultations; to widen channels of communication with other bodies; and to seek opportunities to raise our profile. More than once in this triennium, Standing Committee has discussed whether CUCD is adequately resourced. Although there are, as noted above, several bodies speaking for UK Classics, our remit is distinctive. It seems certain that the task of representing Classics in UK Higher Education will become ever more important and will place increasing demands upon the time, not only of the officers, but of all Standing Committee members. Whether we can carry out this task effectively without asking institutions to increase their subscriptions—allowing us, perhaps, to invest in a permanent secretariat (such as other classical bodies maintain) or a partial buy-out of the Chair's and perhaps other officers' time—is perhaps the most important question that CUCD's membership needs to resolve in the coming months.

GRAHAM SHIPLEY
UNIVERSITY OF LEICESTER
October, 2006

Subject Bodies So Far Registered with the CUCD e-mail list

Association of Art Historians
British Academy (Classics & Ancient History section)
British Association for Slavonic and East European Studies
British Philosophical Association
British Society of Middle Eastern Studies
Classical Association
Council for British Archaeology
Council for College & University English
Council of University Classical Departments
Council of University Deans of Arts, Social Sciences & Humanities
Economic History Society
English Association
Linguistics Association of Great Britain
Media, Communications & Cultural Studies Association
National Association of Music in Higher Education
National Association of Writers in Education
Royal Music Association
Society for the Promotion of Byzantine Studies
Society for the Promotion of Hellenic Studies
Society for the Promotion of Roman Studies
Society of Legal Scholars
Standing Committee of University Departments of Drama
Standing Committee on Modern Greek in the Universities
Standing Council on Dance in Higher Education
Subject Committee for Archaeology
University Council of Modern Languages

All suggestions for additions to this list should be sent to the Secretary or Webmaster.

The Standing Committee of the Council as at September 2006

Elected Members

Chair	Prof. Graham Shipley, University of Leicester
Secretary	Dr Bruce Gibson, University of Liverpool
Treasurer	Dr Patricia Baker, University of Kent
Bulletin Editor	Dr Fiona McHardy, Roehampton University
Statistics	Dr Paul Millett, University of Cambridge
Elections	Dr Costas Panayotakis, University of Glasgow
	Prof. Philip Hardie, University of Oxford
	Dr John Morgan, University of Wales Swansea
	Dr Susanna Phillippo, University of Newcastle upon Tyne
	Prof. Jonathan Powell, Royal Holloway, University of London
	Dr John Rich, University of Nottingham

Co-opted Members

Friends of Classics	Dr Peter Jones, Emeritus, University of Newcastle upon Tyne
Webmaster	Dr Nick Lowe, Royal Holloway, University of London
JACT	Mrs Gill Partington, Our Lady of Sion School, Worthing
Languages	Dr James Robson, Open University
Subject Centre	Prof. Christopher Rowe, University of Durham
	Dr Miriam Plantinga, University of Wales Lampeter

Ex Officio

	Prof. Mike Edwards, Director, ICS

DOING THE SUMS:

A QUANTITATIVE REPLY TO THE AHRC SUPPORTED ERIH PROJECT (CLASSICAL STUDIES)

'We rely upon the rigorous examination of applications by peer
reviewers to ensure that we support world-class research […]'
AHRC Draft Strategy for 2007–2012, Principle 4

Over the last five years, the European Science Foundation (ESF) has been involved in the planning and developing of a citation index for the humanities.[1] This project has so far led to the ongoing development of a European Reference Index for the Humanities (ERIH). The details of the project can be gleaned from the website of the European Science Foundation.[2] Amongst others, the subject area of Classics (there called the discipline of Classical Studies) is scheduled to be covered by the project. More recently, the Arts and Humanities Research Council (AHRC), together with the Higher Education Founding Council for England (HEFCE), has overtly supported the project, in their own more general engagement to facilitate the work carried out by panels that allocate research funding in the future, including the process of the Research Assessment Exercise (RAE). The AHRC and HEFCE now state plainly that after 2008, when the next RAE is scheduled to take place, future research assessment exercises will be carried out by proxy, i.e. on the basis of previously established metrics rather than on actual assessment of the work submitted for the exercise.[3] In the foreword to the consultation document available from the HEFCE website at the above address, Bill Rammell, the Minister of State for Higher Education and Lifelong Learning, states that '[…] the 2008 RAE should proceed largely as planned in parallel with a shadow metrics exercise and that after 2008 it should be replaced with a new and lighter-touch system based largely on metrics', because 'the principle of using information that is already collected routinely to assess research quality and allocate funding must be the right one.'

For this purpose, the AHRC has created its own working party, chaired by Prof. Michael Worton of UCL, to establish how such metrics could be developed in the Humanities. The working party is encouraged to investigate ways of doing so *other* than by the kind of citation index investigated by the ESF. To date, the working party has not announced any results to the wider scholarly community, but has released some first ideas to the RAE 2008 panel members for consultation. The first results are scheduled to be released by mid-October 2006, and it will be the responsibility of the wider scholarly community to subject them to sound and probing scrutiny. Ahead of the working party in its schedule, the ERIH project has in the meantime provoked numerous replies from different subject areas, summaries of which, created by the AHRC, can be viewed on the AHRC website.[4] These summaries fail to disguise the overall concern expressed over the project, as such as well as over the way consultation has been carried out, and one may wish to remain doubtful whether the consultation process to be launched by the AHRC on the results of the working party will be more congenial.

While discussion of the kinds of metrics to be suggested by the AHRC working party will be a matter of the (near) future, the kind of metric explored by the ESF through the ERIH, i.e. the establishment of a citation index, can be—and has already been—subjected to due criticism. It follows a very simple practice: it aims at establishing the quality of a piece of work by the number of active uses of this piece of work by the scholarly community assessed through the frequency of citation; the more citations, the higher the ranking of the work. Although this approach is not in itself unproblematic (Oswald 2006), it is in principle accepted, and metrical assessment of the quality of scholarly output in the sciences is largely dependent on such quantitative analysis. The ERIH project, on the contrary, does not propose to use quantitative

---

[2] www.esf.org.
[3] www.hefce.ac.uk/hefce/2006/metrics.htm.

[4] www.ahrc.ac.uk.

analysis for the establishment of its reference index. Instead, the process is entirely based on non-quantitative methods. This is evident, for instance, in the description of the journal categories that the ERIH project employs (see Appendix 1). Three categories are proposed, A, B and C, of which A should include, '*high-ranking* international publications with a *very strong reputation* among researchers of the field in different countries, *regularly cited* all over the world'; B should include '*standard* international publications with a *good reputation* among researchers of the field in different countries'; and C should include 'research journals with an *important* local/regional *significance* in Europe, *occasionally cited* outside the publishing country though their main target group is the domestic academic community' (my emphasis). The further descriptors offered to rank journals in the correct category state, for instance, that an international journal is defined by '*consistently high-quality* scholarly content' and by a '*broad consensus* within the field concerning international *status* and *visibility*'. The quantitative terms used in these descriptors are not assessed through quantification; rather, it is by way of reputation and impression that journals will be allocated to one or other group by the relevant panels, supported through consultation of the practitioners. Once the journals have been categorised successfully, the project is intended to include monographs and edited volumes in the fullness of time too.

Whilst a qualitative response by those who are critical towards a quantitative assessment of scholarly work is intelligible, it is difficult to accept a purely qualitative approach in the case of those who propose metrical assessment to be the answer to the question of how to assess scholarly work. This reluctance to employ quantitative methods is even more puzzling in the context of a science-oriented model that is based on quantitative analysis. It is the aim of this paper to investigate the potential of metrical assessment of scholarly work in the humanities, by means of the kind of citation index explored by the ESF. By doing so, I wish to establish a working hypothesis as a basis for further quantitative analysis and qualitative discussion, not least for the critical (quantitative) discussion that should ensue from the metrical proposals to be released by the AHRC working party later in the autumn. The hypothesis presented here will of course need testing against a much larger body

of evidence than the evidence that constitutes the sample employed here. To anticipate, the results of the present exercise are largely negative regarding the feasibility of a research assessment exercise based on metrical assessment in the humanities, and thus support the qualitative criticisms made by various subject areas and individuals vis-à-vis the usefulness of the ERIH project briefly mentioned above. The paper adds to these criticisms by demonstrating that research in the humanities can be shown also from a quantitative point of view to follow different rules from research in the sciences, and that, therefore, application of a science-based model of metrical assessment of scholarly work is a futile enterprise in the humanities. In proposing the results I do, I also wish to question why quantitative analysis for the ERIH project has been avoided so far.

Doing the sums

The analysis carried out here is very simple. It has two primary aims: first, to establish the range of journals used by practitioners in Classics for their research, and to assess the quantitative relationship between the various journals used; and second, to establish to what extent the choices made by the sample group are identical with the choices made by the AHRC supported ERIH project. It is not my aim here to create a sample-citation index. Rather, I wish to investigate to what extent the work consulted by practicing Classicists is clustered in a distinct group of journals. It is clear from the purpose of the ERIH, and the AHRC efforts to create a metric-based assessment exercise, that a scholar's accreditation in research assessment exercises would in the future depend entirely to what extent he or she publishes work in journals categorised at the highest level. In practice, the ERIH proposes four categories of journals, A to C, plus journals that are not regarded as making the grade at all and that therefore remain unclassified within the proposed categorisation. Only work published in A-rated journals would achieve maximum accreditation for the scholar (and thus for the scholar's institution). My question therefore is: is there a distinct group of journals that produces regularly cited work more than others, and if so, to what extent does this group of journals dominate the best research? (Only a positive answer to this question justifies a metrics-based research assessment exercise in

the humanities in place of autopsy).

The data employed here come from 17 articles from the recent volume of a UK Classics journal, *The Classical Quarterly*, Vol. 56.1 (2006), published for the Classical Association. This journal has been chosen as it publishes over a wide range, including work on history, language and literature (although less so on material culture), as well as on both Greek and Roman themes. The secondary literature consulted by the article authors has been compiled and is presented in three groups in Chart 1 below. The chart shows the quantitative relationship between the number of contacts made by the authors to i) journal articles, ii) contributions to edited volumes (including conference proceedings, Festschriften, etc), and iii) monographs (including commentaries, editions, etc.). A contact is defined as a reference to a piece of work in any given study; multiple references to the same piece of work in any given study are counted as one (1) contact.

Chart 1

■ *Journals*	227
□ *Edited volumes*	151
▨ *Monographs*	485
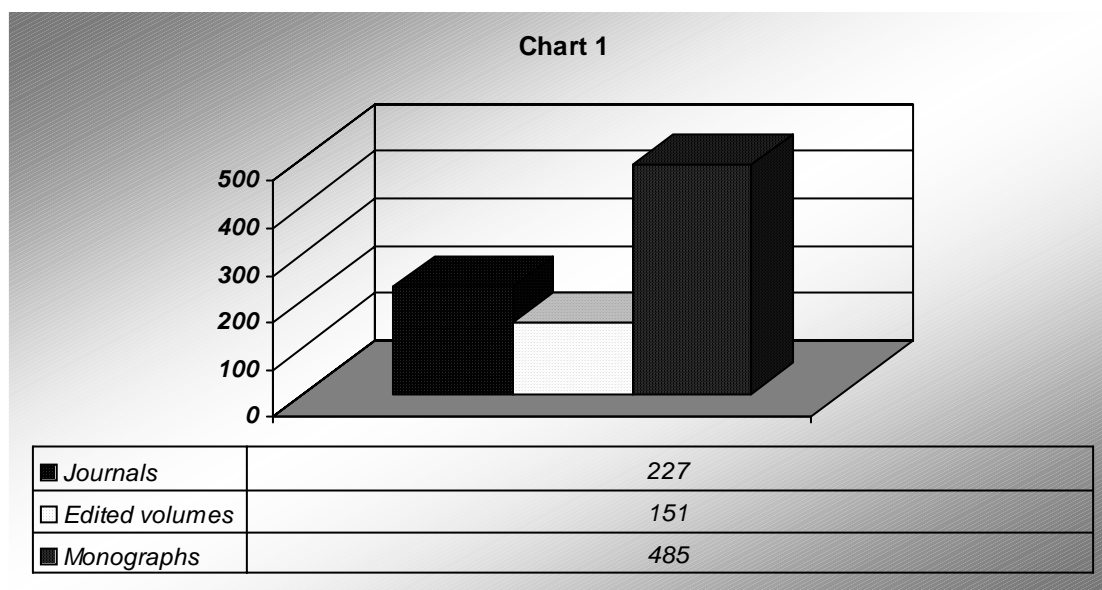
The results are unsurprising: the monograph is the single most important group consulted by the *CQ* authors. Journal articles and contributions to edited volumes together account for 378 contacts, still more than 100 contacts less than the 485 contacts made to monographs. It is also notable here that journals do not dominate the publication of shorter pieces: contributions to edited volumes make up 40% of all short studies referred to.

If furthermore, we graded the *journals* consulted by the *CQ* authors according to the currently proposed ERIH classification (see Appendices 1, 2 and 3), the number of journals consulted that fall into the A category is 32 vis-à-vis 26 in the B category, 5 in the C category, and 28 unclassified journals (Chart 2). In other words, A-rated journals represent only 35% of all classified journals consulted, and are almost matched in number by the journals of categories B and C taken together (which account for 31 journals). If we take all non-A rated journals together, a total of 59 journals, this represents 65% of all

journals consulted, i.e. virtually twice the number of A-rated journals. Going by the number of journals then, A category publication does not appear the dominant feature in Classics journal publication. The CQ authors thought much work in many non-A rated journals important to comment on.

The number of contacts, too, questions whether such dominance exists at all. Of 378 contacts made to different short studies, only 125 were made to publications in A-rated journals, i.e. roughly a third of all short studies referred to. These figures need further qualification: journals classified in the A category are likely to benefit from a higher level of accessibility to researchers, being widely internationally available, often both in traditional paper and in contemporary electronic form, than journals that have been classed differently. (This point too requires sound quantitative analysis to qualify the number of contacts made to journals that benefit from higher accessibility, and thus to qualify the relationship between contacts made and indica-

tion for quality of work assessed.) It seems, Classicists do not cluster their research—in terms of both publication and consultation— primarily in a few select journals, but search wide and in between in their intellectual quest for answers.

Chart 2

■ A	32
□ B & C	31
■ unclassified	28
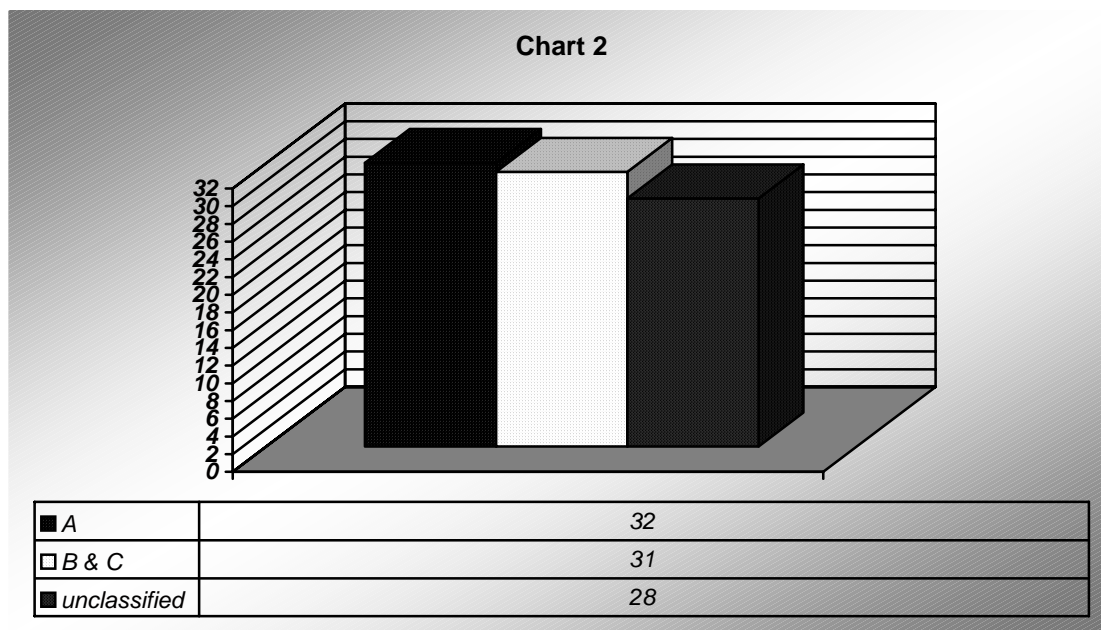
Assuming that A rating will be essential for a scholar and his or her institution to achieve maximum research accreditation in the proposed post-RAE 2008 world, the margin of error would be considerable should accreditation be based on the ERIH or similar metrical approaches. For if we take the choices of the *CQ* authors as our sample for what these authors (who are a much larger group of specialists than the group constituting the Classical Studies panel in the ERIH exercise) considered worthwhile to comment on, the discrepancy becomes immediately apparent: only 55% of the contacts made by them to journal articles fall into the A category; 45% of journal contacts that the *CQ* authors thought important to make fall outside the A category. Publications in the latter however would receive less research accreditation than the former because of the place of publication. If we trust the verdict made by the *CQ* authors in their choice of work consulted, and compare it with the currently proposed ERIH journal classification, accreditation of work based on this classification would get it wrong in almost every second case. It is not clear why a margin of error of nearly 50% should be deemed acceptable in the humanities.

The currently proposed journal classification list is of course still in flux, and changes to it would result in changes here too. But the number of proposed A-rated journals in Classics is already enormous in this list. Of a total of 243 journals, 70 journals, or 29%, are A-rated, despite the ERIH recommendation of 5–20% for A categorisation (see Appendix 1). As we have seen, this already sizeable group of A-rated journals fails to take account of important work published in journals currently not proposed for this category. If on the other hand, we paid once more attention to the *CQ* authors, their natural classification in journals that are heavily cited and journals that are less often cited proposes, from a quantitative point of view, a much stricter categorisation for a top-journal group (see Appendix 2). 70% or 63 journals (of the total of 91 journals cited by the *CQ* authors) show less than the mean number of contacts of 2.5 made by the *CQ* authors. The mean number of contacts of the remaining 28 journals is 4.4, which is achieved by 13 journals only, i.e. by less than 15% of all journals used by the *CQ* authors. More interestingly, perhaps, is the fact that only 11 of the proposed 70 A-rated journals enter the top group here, accounting for just 4.5% of all journals classified in the ERIH Classical Studies list (and again for only just over 15% of all A-rated journals). And only 18 A-rated journals in total make it over the 2.5 hurdle of mean contacts, accounting for 7.4% of all classified journals only (or for just under 26% of all A-

rated journals). These figures suggest that if a purely quantitative approach was employed to create a three-tier grading system of journals in which scholarly significance is measured by number of contacts, it appears the top category should be reserved for a smaller section of all classified journals than the 29% currently employed by the ERIH for Classics. If testing of a larger sample was to confirm the ratios suggested here, this would mean that the currently proposed ERIH Classical Studies list would wrongly allocate the highest accreditation in more than 84% of all cases (i.e. journals) to which at present the highest rating has been credited.

If the ESF, the AHRC and HEFCE want to *measure* research output in the future, they must use quantitative methods to derive a consistent strategy, just as the sciences do. If they wish to renounce on quantification, they need to explain why this should be so, and why they propose to diverge from the kinds of classifications suggested by quantification, as well as how accreditation and discreditation of a journal's value on impressionistic terms (and without autopsy of individual contributions) can be at all justified. The *CQ* authors who do evidently not base their evaluation of research on metrical assessment but on autopsy, appear themselves not so foolish as to disregard or discredit a vast bulk of journals that could, from a quantitative point of view, not make the top-grade, even if we allowed for a generous A-rating of nearly a third of all classified journals as the ERIH Classical Studies list currently does.

Naturally, the *CQ* authors may have got it wrong on occasion, i.e. they may have allocated too much significance to some work (or indeed too little to others); but this would apply to all work consulted by them, and whatever bias is inherent in their choices would affect work published in journals that the ERIH Classics panel classified as A, B, C or not at all. It seems also evident from the analysis carried out here that, unsurprisingly, scholars tend to publish work in journals that publish work important to their research, a commonplace one may think, but an important one in any attempt to analyse the scholarly significance of any given journal. For to understand fully the prevalence of contacts made to contributions in *The Classical Quarterly* by its authors (19 contacts in total, see Appendix 2), whilst showing the importance of

work published in this journal, we need to take into account that the authors may have chosen *CQ* as a place of publication because of the publication of work in *CQ* that is important to them. If, on the other hand, we take a look at the article that is responsible for 5 out of 12 contacts to *The Journal of Roman Studies*, namely Serrati's piece on the treaties between Rome and Carthage, it becomes immediately obvious that the choice of contacts made, like the theme of the article, would recommend the article for publication in *JRS*. In other words, work on any given theme or topic attracts discussion of work published in journals that provide central stage for that theme or topic. Hence, to gain a full picture of the importance of journals for all types of work carried out under the Classics umbrella, a large sample containing journals covering all aspects of Classical studies need to be analysed accordingly. It does not follow from this observation that journals that specialise more than others in their choice of themes covered by them, be it geographically, chronologically or thematically, should by default be deemed of a lesser scholarly quality; quite the opposite is suggested by the figures derived from the analysis carried out here.

Conclusions

It may be worthwhile to sum up, and reiterate, some of the points made above. The *CQ* authors evidently valued and thought important for their work short studies published outside journal category A. In fact, it may need emphasising that the vast majority of contacts fell outside the top-journal grade. In view of the wider context in which the ERIH is to be placed, i.e. the attempt to develop a form of research output assessment through metrics, some interesting questions have been raised which need to be taken further.

First, the location of publication of the studies to which the majority of contacts in our sample make reference would under the proposed scheme be less suitable to further the authors' careers (and the research income of their institutions), than those studies to which the lesser number of contacts makes reference. In other words, the judgement made by the *CQ* authors, taken as a hypothetical model for the pattern of a journal citation index in Classics (not an article citation index), is not compatible with the proposed ERIH journal ranking in Classics, for it shows that a vast range of journals publishes

work (many more than what could possibly be squeezed into the recommended 5–20% target range of category A journals) that is noteworthy and that requires active comment and debate by the scholarly community. In other words, the classification in A, B, C (and implicitly non-graded) journals, runs, on the basis of the evidence analysed here, counter the dynamics typically associated with a citation index as used in the sciences. There, publication in a top-journal is, ideally, more often associated with a high citation count than publication in a medium or lower grade journal. But *CQ* 56.1 (2006) does not fit this model: the contributors evidently do not view journal location as an indicator for work that is noteworthy, but made their choices largely independently of publication locations. The journals currently classed as A-category journals, fail, on current analysis, to outdo (and indeed to match), the number of contacts made to studies in non-category A journals; they do not dominate the field.

Indeed, one of the interesting results of the present analysis is that the mean number of contacts per journal is not even twice as high for category A journals as for category B journals, and still only just over two times as high for A category journals compared with all journals classified outside the A category—and this has not even taken into account the fact that journals classified in the A category appear to have on average a larger number of articles published in them per volume (and indeed often more volume numbers per year) than journals with a lower or no classification, and hence a greater chance of contacts per journal. (E.g., *The Classical Quarterly*, which appears twice a year, and which is A-rated, published 39 full-lengths articles in the last two issues [56.1 and 55.2]; *Classics Ireland*, which appears once a year, and which is C-rated, published three full-length articles in 2006.) To put it differently: it would be worthwhile to quantify the relationship between number of contacts per journal and average number of articles per journal and year to derive a mean number of likely contacts per article. On present analysis, it seems unlikely that category A journals will at all achieve higher results in this than journals outside the A cate-gory—which must be viewed against the clear differences of article citation likelihood in the sciences depending on journal standing (Oswald 2006 suggests a mean number of citations per article that is roughly four times higher in a journal of the most prestigious standing than even in journals of good to medium standing). *CQ* 56.1 (2006), then, suggests that the sciences and the humanities do not work according to the same rules of scholarship. It suggests further-more that to propose a metrical analysis of research output in the humanities that is modelled on schemes developed in (and that are helpful for) the sciences is intrinsically flawed. It should not occasion surprise though that scholars largely unfamiliar with quantitative methods should have missed such an essential point in their quest to support the ERIH and metric-based approaches to assess the quality of work carried out in the humanities. In short, rigorous examination of applications and appli-cants, research assessment contributions and contributors, a principle that the AHRC includes in its Draft Strategy for 2007–2012, appears impossible to be carried out by proxy in the humanities.

The example chosen here is isolated. It is an excruciatingly small sample. But I have my doubts that the patterns discerned here would substantially change if the sample size was increased—as indeed it should be. As stated at the outset, the result of the analysis carried out here is meant to function as a working hypothe-sis, and wants to encourage due and sophisti-cated criticism. It is a hypothesis against which to test further evidence. And to test further evidence is a *sine qua non* for any further discussion on the ERIH and metrics-based research assessment exercises in the humanities. The current unqualified, because entirely unquantified discussion at all levels cannot make for an informed analysis of the matter—for it lacks scholarly rigour demonstrated in wide and probing analysis of the evidence. As a minimum, a sound body of journals, covering all aspects of work carried out under the Classics umbrella, and accounting for a sizeable number of journals in the field, needs to be analysed—i.e. quanti-fied—regarding their authors' citation patterns. The minimum period under review should be not less than ten years, and preferably closer to twenty years to cover as much if not all of the period affected by the RAE, and thus to allow for as detailed a conclusion as possible. (The same will need to be carried out for other subject areas too, as whatever pattern Classics will produce cannot be taken for granted for other

areas in the humanities.) Chronological differences need to be studied too, i.e. the extent to which quality and range of the publications of any given journal differs over time. So too geographical differences, i.e. the extent to which a scholar working in, say, an Anglo-American tradition has more or less recourse to certain types of publication than a scholar working in, say, a Germanic tradition, etc. (John Serrati, for instance, makes twelve contacts to seven category A journals—all but one of which are products of Anglo-American scholarship, and makes an equal number of contacts to short studies published in edited volumes, whilst Jonas Grethlein restricts himself to eight contacts to eight different journals—of which only 3 belong to the A category—but makes fourteen contacts to short studies in edited volumes, and 29 contacts to monographs). There may also be differences within Classical scholarship: the Roman historian may be more at home in the world of the A journal than the Greek philologist or vice versa; similarly, there may simply be differences in style and type of referencing from scholar to scholar: a single, unified assessment model may thus not be rational (see also Appendix 4). Again, this question, like the previous ones, cannot be answered through allocation of scholarly importance on the basis of a few individuals' final choices; the answers lie within the remit of quantitative analysis.

Without such quantitative analysis, however, there is no need to return to the boardroom table, and to waste any further time, effort and, most of all, public monies at this stage. Although, as this analysis has shown, there is no way of avoiding what is likely to be a massive use of human time and effort, as well as public monies, to derive a sound quantitative assessment of the evidence—for despite how Rammell summarises the benefits of a metrics-based research assessment exercise, there does not exist 'information that is already collected routinely' in Classics, and other subject areas in the humanities are likely to lack such routinely collected information too. Such collection would need to be specifically carried out now, and the present analysis has pointed to just some of the aspects that would need to be taken care of to make the exercise worthwhile. Personally, I would want to remain doubtful whether such an effort to collect, now, the necessary data and to analyse it quantitatively would at all create any

savings of time and effort as Rammell implies, not to speak of public monies, even in the long run, compared with the monies spent currently on the regular 5-yearly research assessment exercises that are based on autopsy of submitted work. Those 5-yearly research assessment exercises may not be the best way of allocating research funding, but at least they are cheap, comparatively speaking. Again, the humanities may be different in this respect from the sciences, and the former are likely to be less costly to assess on the basis of regular assessment than on that of pre-established metrics. Anyone in favour of a switch to research assessment exercises in the humanities that are based on metrics needs to demonstrate the implied financial gain from such a switch—and not just assume it. And anyone heavily engaged in the development of metrical assessment modes, like the members of the AHRC working group, needs to convince their colleagues in institutions other than their own that their interest and engagement in this exercise is not essentially driven by a strong loyalty towards their own institution and its working methods.

To resume discussion without quantitative analysis, however, means to ignore the evidence, if not purposefully to foreclose the conclusions that quantitative analysis may bring with it. It is, in fact, one of the grotesque implications of the ERIH enterprise that those wishing to count scholarly output by numbers wish to do so precisely without numbers. If they did have recourse to quantitative analysis, they would soon realise their misguided approach—and maybe, this is precisely why they have avoided quantitative analysis so far. The working party set up by the AHRC in collaboration with HEFCE needs to do better than that. It follows that any proposals from that quarter that are lacking quantitative analysis—and a balance sheet—should not be regarded as serious suggestions. We as a scholarly community whose practitioners, by and large, are only marginally confronted with quantitative analysis, should not allow those who think that our thoughts and ideas can reasonably be measured by anything else than the biggest computer on the planet, the human brain, to do so without providing quantitative proof for their currently unsupported assumptions. And don't be afraid to demand such proof: as the above example has

shown, there is nothing to fear in quantifying the Classics, if it is properly done. The sums look promising.

Addendum

Between completion and publication of this piece, the report of the expert group set up jointly by the AHRC and HEFCE, called 'Use of research metrics in the arts and humanities', has been released. Apart from failing to provide any concrete proposals for future research assessment exercises—the group recommends in their place simply a set of key principles and broad operational features which do not offer much that is new or different from the principles and operational features for the RAE 2008—it proposes that 'for the purposes of research assessment, there is no fundamental difference between STEM subjects (science, technology, engineering and mathematics) and arts and humanities disciplines'. As the above exercise should have shown, there is considerable reason to doubt that no differences exist between STEM subjects and arts and humanities subjects in their research practices and the ways research outputs are achieved and can subsequently be assessed. The question is not whether bibliometrics can or cannot be used to assess arts and humanities subjects, but rather to what extent we consider subjects belonging to one or other of the two main groups different from one another regarding research practices and all that goes with that. If the argument presented above gets anywhere then there exists reasonable reason to question the key principle proposed by the AHRC/HEFCE expert group. And if the issue at stake is worth following up, expert groups of all pedigrees should not be allowed to get away with proposing any key principle without supplying evidence for any underlying assumptions; the key principle quoted here lacks any evidentiary support from those who hold it, it is presented as a matter of opinion, and not of fact. Ironically, one of the main conclusions of the AHRC/HEFCE expert group is that '[…] metrics alone will not allow the overall performance and quality of research to be assessed at individual or departmental level'; it appears that one could have reached this conclusion without all the fuss. If anything, the report of the AHRC/HEFCE expert group is in itself evidence of how wasteful are the efforts of those wishing to explore a metrics-based or metrics-supported research assessment exercise, in regard to public monies as well as individual time and effort: is it time to demand of them to justify this waste … maybe by asking for a fully-costed balance-sheet for their exercise?

ULRIKE ROTH
UNIVERSITY OF EDINBURGH

Bibliography

T.D. Barnes, 'An urban prefect and his wife', *CQ* 56.1 (2006), 249–256

P. Bosman, 'Selling cynicism: the pragmatics of Diogenes' comic performances', *CQ* 56.1 (2006), 93–104

E. Carawan, 'Amnesty and accountings for the Thirty', *CQ* 56.1 (2006), 57–76

D. Collins, 'Corinna and mythological innovation', *CQ* 56.1 (2006), 19–32

G.G. Fagan, 'Bathing for health with Celsus and Pliny the Elder', *CQ* 56.1 (2006), 190–207

J. Grethlein, '*Nam quid ea memorem*: the dialectical relation of *res gestae* and *memoria rerum gestarum* in Sallust's *Bellum jugurthinum*', *CQ* 56.1 (2006), 135–148

O. Hekster and J. Rich, 'Octavian and the Thunderbolt: the temple of Apollo Palatinus and Roman traditions of temple building', *CQ* 56.1 (2006), 149–168

N. Kanavou, 'Personal names in the *Vita Aesopi* (*Vita G* or *Perriana*)', *CQ* 56.1 (2006), 208–219

J.T. Katz and K. Volk, 'Erotic hardening and softening in Vergil's eighth *Eclogue*', *CQ* 56.1 (2006), 169–174

V.J. Liapis, 'Achilles Tatius as a reader of Sophocles', *CQ* 56.1 (2006), 220–238

G.W. Most, 'Apollo's last words in Aeschylus' *Eumenides*', *CQ* 56.1 (2006), 12–18

A.J. Oswald, 'An examination of the reliability of prestigious scholarly journals: evidence and implications for decision-makers', (2006 Papers) http://www.andrewoswald.com

C. Patterson, '"Citizen cemeteries" in classical Athens?', *CQ* 56.1 (2006), 48–56

N. Rood, 'Implied vengeance in the simile of grieving vultures (*Odyssey* 16.216–19), *CQ* 56.1 (2006), 1–11

J. Serrati, 'Neptune's altars: the treaties between Rome and Carthage (509–226BC)', *CQ* 56.1 (2006), 113–134

M. Wright, '*Orestes*, a Euripidean sequel', *CQ* 56.1 (2006), 33–47

APPENDIX 1: ERIH Summary Guidelines for Use in the Consultation Process

Coverage

The following disciplines are covered: Anthropology, Archaeology, Art and Art History, Classical Studies, Gender Studies, History, History and Philosophy of Science, Linguistics, Literature, Music and Musicology, Oriental Studies and African Studies, Pedagogy and Educational Research, Philosophy, Psychology, Religious Studies and Theology. Media studies will be covered more fully in a second round of consultation. Journals covering overlapping fields may be found in the lists produced by two or more panels (possibly even with different gradings).

Standards

All journals included must fulfil normal international academic standards, i.e. selection of articles is based on an objective review policy. This quality control is normally through peer review, and it is expected that journals would depart from peer review only where there is another system ensuring quality control. In some scholarly traditions peer-review is an unfamiliar procedure. It is one aim of ERIH to encourage top-journals to adopt a coherent peer-review system.

The journals must fulfil basic publishing standards (i.e. ISSN, timeliness of publication, complete bibliographic information for all cited references, full address information for every author).

It is expected that journals would depart from peer review only where there is another system ensuring quality control. ERIH is sensitive to scholarly traditions in which peer-review is an unfamiliar procedure.

Only journals that fall into the following three categories should be included:

1) Journals category A: i.e. high-ranking international publications with a very strong reputation among researchers of the field in different countries, regularly cited all over the world.

2) Journals category B: i.e. standard international publications with a good reputation among researchers of the field in different countries.

3) Journals category C: research journals with an important local / regional significance in Europe, occasionally cited outside the publishing country though their main target group is the domestic academic community.

Nota bene

- For categories A and B, journals published in the whole world can be considered
- For category C, only European journals must be considered.
- It is recommended that in category A, only 5 to 20% of the total list should appear; this percentage target will differ from one discipline to another.

"International journals"

1) A journal is international (Categories A and B) when the following requirements are fulfilled in addition to those that apply to all journals:

- A genuine, varied and regular international cohort of contributors and readership
- Consistently high-quality scholarly content
- Broad consensus within the field concerning international status and visibility

2) In addition, they will have some, though not necessarily all, of the following characteristics:

- Active international advisory board
- Open to unsolicited contributions
- Highly discriminating and selective in the choice of articles published
- Published on time and to an agreed schedule

The difference between category 'A' and category 'B' journals is likely to be the degree to which they conform to 1) above, and both the number of characteristics under 2) to which they conform as well as the degree of conformity. Generally, 'A' journals should conform to more of these characteristics, and to a greater extent, than 'B' journals.

Language

Main international languages in this context are English, French, German, Spanish and Russian. However, journals in other languages can also be "international", when they are being used as "forum language" for specific research communities. Similarly, there may be non-European journals, that are international in character, in languages other than European languages, which can be included in categories A and B.

Process

- ESF (SCH) Member Organisations provide initial lists of journals in 15 disciplines, and contribute through further consultation
- 15 ERIH Expert Panels analyse, harmonise and finalise lists of graded journals in categories A, B and C.
- ERIH Steering Committee supervises and unifies methodology ("bottom up") and reports to Standing Committee for the Humanities, and through ESF / SCH to HERA / European Commissions
- ESF Standing Committee for the Humanities approves final reference lists and ensures acceptance
- They develop a financial continuity plan and make recommendations on future development of ERIH

APPENDIX 2: Journals cited by the *CQ* authors; number of contacts made; ERIH classification in the ERIH Classical Studies list (see Appendix 1)—provisional

Table 1.1 (in alphabetical order)

Abbreviations:	'Contacts made'	'ERIH classification'
A & A	1	B
AA	1	C
ABSA	2	A
Acta Classica	3	C
AJA	3	A
AJAH	1	u/c
AJP	6	A
Annales de la Faculté des Lettres/SH d'Aix	1	u/c
Annuaire de l'Université de Sophia: Lettres	1	u/c
ΑΠΘ	1	u/c
Arch. Laz.	1	u/c
Archaiognosia	1	C
Arethusa	8	B
ARID	1	B
ASNP	1	B
Athenaeum	5	A
BABesch	1	u/c
BCH	1	A
BICS	3	B
BMCR	1	u/c
Boll. d'Arte	2	u/c
Bull. Soc. Etudes Anciennes Quebec	1	u/c
BZ	1	A
Cahiers de l'Ecole Normale Supérieure	1	u/c
Classical Antiquity	2	A
Colby Quarterly	1	u/c
CP	5	A
CQ	19	A
CR	3	u/c
Dial. Arch.	2	u/c
Die Antike	1	u/c
Elenchos	1	B
Emerita	2	B
EOS	2	C
Eranos	4	B
G & R	5	B
Glotta	1	A
Gnomon	1	u/c
GRBS	3	A
Gymnasium	1	B
Hecuba	1	u/c
Helios	2	B
Hermes	6	A
Hesperia	1	A
Historia	3	B
Horos	1	B
HSCP	6	A
Humanitas	1	C
ICS	4	B
JDAI	1	A
JHS	7	A

Journal	Contacts	Grade
JRA	1	A
JRS	12	A
Klio	2	B
Kokalos	1	B
La Nouvelle Clio	1	u/c
Latomus	2	B
MAAR	1	u/c
Maia	2	B
MD	1	A
MDAI(R)	3	A
MEFRA	7	A
MH	3	A
Mnemos.	2	A
MSL	1	u/c
NJb	1	u/c
Numismatic Chronicle	1	u/c
PACA	1	u/c
PBSR	3	A
PCPS	2	A
Phoenix	3	B
PP	1	A
Proc. Dan. Inst. Ath.	1	u/c
QS	1	B
Ramus	1	B
RCCM	1	u/c
REA	1	A
REG	1	A
Rend. Ist. Lomb.	1	u/c
Rend. Nap.	1	u/c
Rend. Pont. Acc.	4	u/c
RhM	7	A
RIDA	1	B
SDHI	1	B
SIFC	4	B
Syllecta Classica	1	u/c
TAPA	7	A
Traditio	2	B
Wiener Studien	2	B
ZPE	4	A
ZRG	1	A

TOTAL:					
TOTAL:	91 journals	100%	227 contacts	100%	32A 26B 5C 28u/c
	49 journals	53.8%	1 each = 49	21.6%	11A 11B 3C 24u/c
	14 journals	15.4%	2 each = 28	12.3%	4A 7B 1C 2u/c
	10 journals	11.0%	3 each = 30	13.2%	5A 3B 1C 1u/c
	05 journals	05.5%	4 each = 20	08.8%	1A 3B 0C 1u/c
	03 journals	03.3%	5 each = 15	06.6%	2A 1B 0C 0u/c
	03 journals	03.3%	6 each = 18	07.9%	3A 0B 0C 0u/c
	04 journals	04.4%	7 each = 28	12.3%	4A 0B 0C 0u/c
	01 journal	01.1%	8 each = 08	03.5%	0B 1B 0C 0u/c
	01 journal	01.1%	12 each = 12	05.3%	1A 0B 0C 0u/c
	01 journal	01.1%	19 each = 19	08.4%	1A 0B 0C 0u/c

Table 1.2 (in ascending order of number of contacts)

Abbreviations:	'Contacts made'	'ERIH classification'
A & A	1	B
AA	1	C
AJAH	1	u/c
Annales de la Faculté des Lettres/SH d'Aix	1	u/c
Annuaire de l'Université de Sophia: Lettres	1	u/c
ΑΠΘ	1	u/c
Arch. Laz.	1	u/c
Archaiognosia	1	C
ARID	1	B
ASNP	1	B
BABesch	1	u/c
BCH	1	A
BMCR	1	u/c
Bull. Soc. Etudes Anciennes Quebec	1	u/c
BZ	1	A
Cahiers de l'Ecole Normale Supérieure	1	u/c
Colby Quarterly	1	u/c
Die Antike	1	u/c
Elenchos	1	B
Glotta	1	A
Gnomon	1	u/c
Gymnasium	1	B
Hecuba	1	u/c
Hesperia	1	A
Horos	1	B
Humanitas	1	C
JDAI	1	A
JRA	1	A
Kokalos	1	B
La Nouvelle Clio	1	u/c
MAAR	1	u/c
MD	1	A
MSL	1	u/c
NJb	1	u/c
Numismatic Chronicle	1	u/c
PACA	1	u/c
PP	1	A
Proc. Dan. Inst. Ath.	1	u/c
QS	1	B
Ramus	1	B
RCCM	1	u/c
REA	1	A
REG	1	A
Rend. Ist. Lomb.	1	u/c
Rend. Nap.	1	u/c
RIDA	1	B
SDHI	1	B
Syllecta Classica	1	u/c
ZRG	1	A
ABSA	2	A
Boll. d'Arte	2	u/c
Classical Antiquity	2	A
Dial. Arch.	2	u/c
Emerita	2	B
EOS	2	C
Helios	2	B
Klio	2	B
Latomus	2	B
Maia	2	B
Mnemos.	2	A

PCPS	2	A
Traditio	2	B
Wiener Studien	2	B
Acta Classica	3	C
AJA	3	A
BICS	3	B
CR	3	u/c
GRBS	3	A
Historia	3	B
MDAI(R)	3	A
MH	3	A
PBSR	3	A
Phoenix	3	B
Eranos	4	B
ICS	4	B
Rend. Pont. Acc.	4	u/c
SIFC	4	B
ZPE	4	A
Athenaeum	5	A
CP	5	A
G & R	5	B
AJP	6	A
Hermes	6	A
HSCP	6	A
JHS	7	A
MEFRA	7	A
RhM	7	A
TAPA	7	A
Arethusa	8	B
JRS	12	A
CQ	19	A

Journals classified in other subject areas:

Arcadia	1
Archaeology	1
Archive für Geschichte und Philosophie	1
Art Forum	1
Cambridge Archaeological Journal	1
Epet. Byz.	1
History of Religions	1
IEJ	1
Jahresb. der schweiz.Geisteswissensch. Gesell.	1
Journal of Medit. Archaeology	1
Journal of Medit. Studies	1
London Review of Books	1
Man	1
Monuments Piot	1
Paideia	1
Recherches Augustiniennes	1
Revue des études augustiniennes	1
Revue Philosophique	1
Rhetoric Review	1
RSI	1
Semiotica	1
Thracia	1
Tulsa Studies in Women's Literature	1
World Archaeology	1
TOTAL	24

APPENDIX 3: Number of contacts made by the *CQ* 56.1 (2006) authors to journal articles, contributions to edited volumes and monographs by article.

Scholar		Journals	Ed. Vols.	Monographs
Rood	TOTAL:	19	3	21
	AJP	1		
	Cahiers de l'Ecole Normale Sup.	1		
	CP	2		
	CQ	3		
	G & R	3		
	Helios	1		
	History of Religions	1		
	HSCP	2		
	JHS	3		
	TAPA	2		
Most	TOTAL:	2	1	21
	NJb	1		
	TAPA	1		
Collins	TOTAL:	23	6	17
	Acta Classica	1		
	AJP	1		
	Annales de la Faculte L/SH d'Aix	1		
	Arethusa	2		
	BCH	1		
	Colby Quarterly	1		
	CQ	3		
	CR	1		
	EOS	2		
	Hermes	1		
	SIFC	1		
	Syllecta Classica	1		
	TAPA	1		
	Tulsa Studies in Women's Literature	1		
	ZPE	2		
Wright	TOTAL:	27	4	35
	AJP	1		
	Arethusa	3		
	BICS	1		
	Classical Antiquity	1		
	CQ	1		
	Eranos	1		
	Hermes	2		
	HSCP	3		
	ICS	3		
	London Review of Books	1		
	Maia	1		
	PCPS	1		
	Ramus	1		
	RhM	1		
	SIFC	1		
	TAPA	2		
	Traditio	2		
	Wiener Studien	1		

Patterson	TOTAL:	12	7	16
	ABSA	1		
	Archaeology	1		
	Cambridge Archaeological Journal	1		
	CQ	1		
	CR	1		
	Gnomon	1		
	Helios	1		
	Horos	1		
	JHS	1		
	Journal of Mediterranean Archaeology	1		
	Journal of Mediterranean Studies	1		
	World Archaeology	1		
Caravan	TOTAL:	11	3	28
	CQ	2		
	Hesperia	1		
	JHS	1		
	Klio	1		
	MH	1		
	Mnemosyne	1		
	REG	1		
	RhM	1		
	RIDA	1		
	ZRG	1		
Bosman	TOTAL:	11	19	15
	Elenchos	1		
	Rhetoric Review	1		
	JHS	1		
	Arcadia	1		
	Archiv für Geschichte der Philosophie	1		
	Art Forum	1		
	BICS	1		
	CR	1		
	Man	1		
	Semiotica	1		
	Revue Philosophique	1		
Serrati	TOTAL:	25	13	50
	Acta Classica	1		
	AJA	1		
	AJAH	1		
	AJP	1		
	ARID	1		
	Bull. Soc.Etudes Anciennes du Quebec	1		
	CP	1		
	CQ	2		
	Emerita	1		
	G & R	1		
	Historia	1		
	Humanitas	1		
	JRS	5		
	Kokalos	1		
	Latomus	1		
	MSL	1		
	PBSR	1		
	Phoenix	1		

	REA	1		
	RSI	1		
	SDHI	1		
Grethlein	TOTAL:	8	14	29
	A & A	1		
	CA	1		
	G & R	1		
	Gymnasium	1		
	JRS	1		
	Maia	1		
	Mnemosyne	1		
	PACA	1		
Hekster/Rich	TOTAL:	50	20	58
	AA	1		
	AJA	1		
	AJP	1		
	Arch. Laz.	1		
	Athenaeum	3		
	BABesch	1		
	BICS	1		
	BMCR	1		
	Boll. d'Arte	2		
	CQ	2		
	Dial. Arch.	2		
	Emerita	1		
	Hermes	1		
	Historia	1		
	JDAI	1		
	JRA	1		
	JRS	2		
	La Nouvelle Clio	1		
	MAAR	1		
	MDAI(R)	3		
	MEFRA	7		
	MH	2		
	Monuments Piot	1		
	Numismatic Chronicle	1		
	PBSR	2		
	Proc. Dan. Inst. Athens	1		
	RCCM	1		
	Rend. Ist. Lomb.	1		
	Rend. Pont. Acc.	4		
	RhM	1		
	ZPE	1		
Katz/Volk	TOTAL:	6	5	19
	Arethusa	1		
	CPh	1		
	HSPh	1		
	Phoenix	1		
	RhM	1		
	SIFC	1		
Woodman	TOTAL:	11	4	34
	Athenaeum	1		
	CQ	1		
	Hermes	3		
	Historia	1		

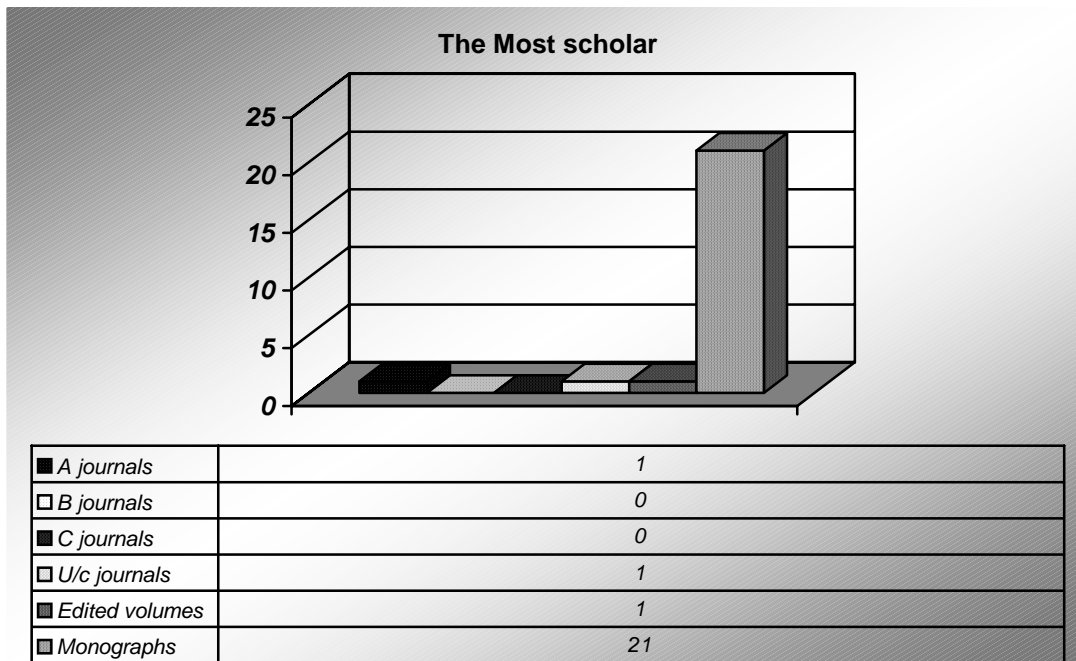|				
	JRS	3		
	Latomus	1		
	PCPS	1		
Fagan	TOTAL:	13	16	23
	AJA	1		
	ASNP	1		
	Die Antike	1		
	Epet. Byz.	1		
	IEJ	1		
	JHS	1		
	JRS	1		
	Klio	1		
	PBSR	1		
	PP	1		
	Rend. Nap.	1		
	RhM	1		
	ZPE	1		
Kanavou	TOTAL:	7	16	39
	Athenaeum	1		
	CQ	1		
	TAPA	1		
	RhM	1		
	BZ	1		
	ICS	1		
	ΑΠΘ	1		
Liapis	TOTAL:	21	12	49
	ABSA	1		
	Acta Classica	1		
	AJP	1		
	Annuaire de l'Uni. Sophia: Lettres	1		
	Archaiognosia	1		
	Arethusa	2		
	CP	1		
	CQ	1		
	Glotta	1		
	GRBS	3		
	Hecuba	1		
	JSGG	1		
	MD	1		
	Paideia	1		
	QS	1		
	RhM	1		
	SIFC	1		
	Thracia	1		
Barnes	TOTAL:	6	8	31
	CQ	2		
	Phoenix	1		
	Recherches Augustiniennes	1		
	Revue des études augustiniennes	1		
	Wiener Studien	1		

Note: Excluded from the analysis are the articles by Barnett, Johnson, Konstan and Ramelli, and Wolfsdorf because their references fall largely outside the remit of classical studies, as well as all shorter notes published in *CQ* 56.1 (2006).
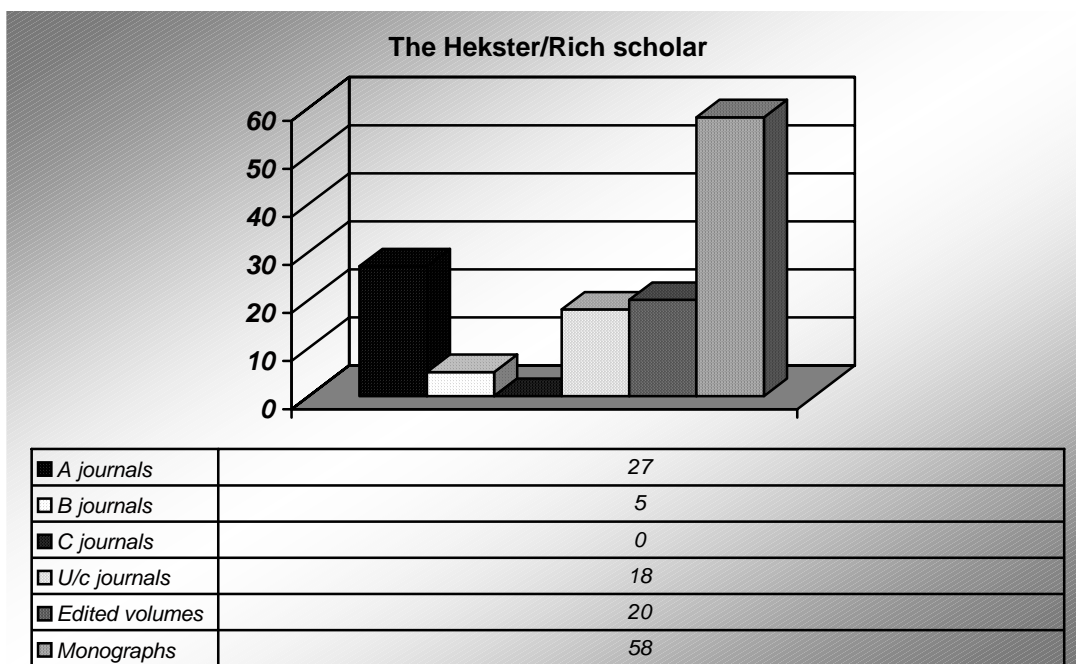
To test the referencing pattern of your own scholarship, collate all references to journal articles, contributions to edited volumes and monographs in one of your own pieces of work and match them against the three sample types described and graphically laid out below. Then consult the key below to learn about the type of scholarship you subscribe to.

Can't find a match? Create your own profile and compare it with those of your colleagues. Happy compiling!
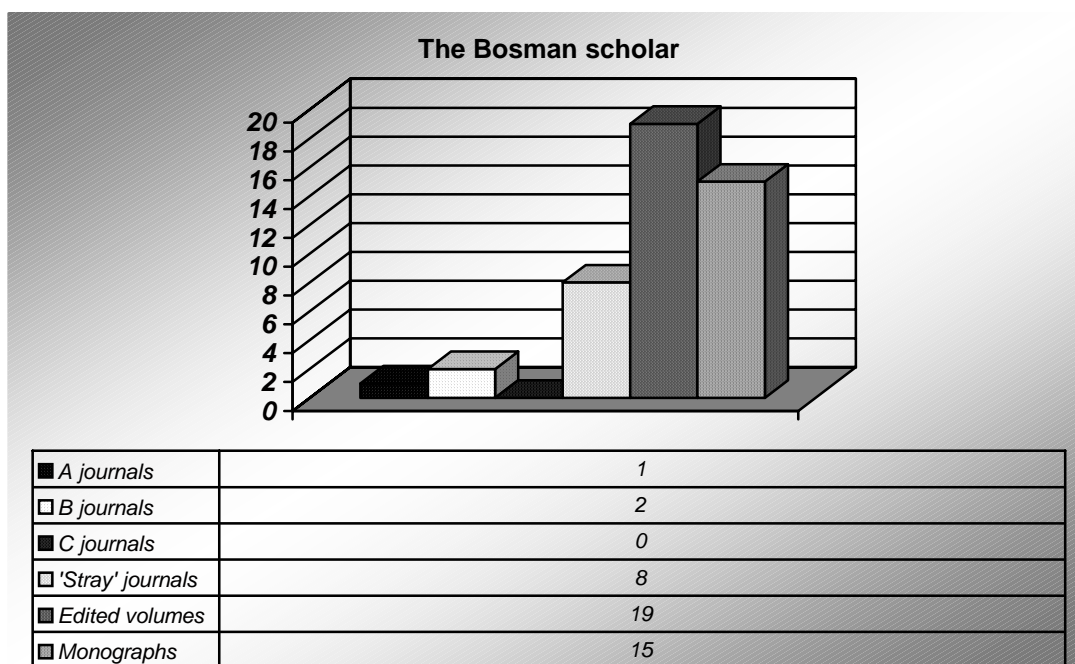
Type 1

The Most scholar

■ A journals	1
□ B journals	0
■ C journals	0
□ U/c journals	1
■ Edited volumes	1
▨ Monographs	21

Type 2

The Hekster/Rich scholar

■ A journals	27
□ B journals	5
■ C journals	0
□ U/c journals	18
▨ Edited volumes	20
▨ Monographs	58

Type 3

The Bosman scholar

■ *A journals*	1
☐ *B journals*	2
■ *C journals*	0
☐ *'Stray' journals*	8
■ *Edited volumes*	19
☐ *Monographs*	15

KEY

Type 1: The Most scholar—*Why wander?*
You do not worry about the recent journal editions, or indeed about journals as a whole. Instead you focus on the great work produced by your colleagues in their monumental monographs, and you will no doubt return, after a short flirt with the journal world, to the world of monographic achievements yourself.

Type 2: The Hekster/Rich scholar—*The importance of being earnest*
You read extensively and intensively, and are at home both in the world of the short study as well as in that of the monograph. You do not even refrain from consulting Danish journals, and show generally a continental affiliation. You also make sure that your wide reading is documented in your footnotes, as is your own work on the topic.

Type 3: The Bosman scholar—*The 'stray' cat*
You push your reading to extremes in as much as you do not show much of a bias for or against a certain type of publication. But you are also an *enfant terrible* of the classical world, because much of the scholarship that you consult cannot neatly be classified in a Classical Studies list.

ERIH: New Directions and Future Developments

This response to 'Doing the sums …' aims to update the classics community about the state of work of ERIH, and about current debates regarding the future development of the project. At the end of the response, an addendum provides some corrections to the picture painted in 'Doing the sums …' to the way in which the recently established HEFCE/AHRC Expert Group on Research Metrics, chaired by Prof. Michael Worton, and ERIH are related.

I have been involved with ERIH since August 2005, first as Chair of the Archaeology Expert Panel, then as Chair of the Panel Chairs Group. It seems to me that the issues surrounding the use and value of 'metrics' in assessing research, including humanities research, are complex—this is one reason why I am involved in ERIH. The simple quantitative 'metrics' advocated by the author of 'Doing the sums …' muddies the waters in these debates.

ERIH is a project coordinated by the European Science Foundation (ESF) on behalf of its member organisations, which include research councils, such as AHRC, but also organisations delivering research, such as the French CNRS, and academies. ERIH is funded by those organisations and by the European Commission under the ERA-Net project HERA (Humanities in the European Research Area). ESF, the HERA-network, and ERIH itself, all aim at strengthening humanities research in the context of scientific research in Europe, by making more visible the quality of scientific output produced by humanities researchers. While ERIH focuses currently on journals, coverage of conference contributions and monographs is envisaged for a later stage. ERIH is governed by a Steering Committee of academics, chaired by Prof. Alain Peyraube, member of the ERC Scientific Council.

ERIH aims at the production of categorised lists of good research journals in—currently—fifteen domains of the humanities. With its categorization of A/B/C, ERIH wishes to highlight the different audiences (international and national) to which high-quality research can be addressed. In order to take into account of the very uneven accessibility of research output from different areas of Europe (due to resources, distribution, language), ERIH adopted from the beginning an approach based on peer review.

ERIH was initially conceived of as the basis for a 'citation index'. The project in fact used to be called ECIH. It was geared towards strengthening European humanities compared to their poor coverage in existing citation indices. The need for this is particularly felt by humanities research councils and funding bodies across Europe, who are trying to access a fair share of national research budgets for humanities subjects and find themselves in competition with 'big science' for funds. At present it is difficult to provide a convincing demonstration of the great strength of European humanities research to scientists. It soon became clear, however, that such an enterprise entails a great deal of preparatory work, and that more subtle measurement tools are needed to do justice to the complex range and types of research output in the humanities. But, if we as academics do not engage with this process, it will happen one way or another, without us.

The Expert Panels are envisaged as the peer-review base for ERIH. As in every peer-review process, the scholarly debate on quality is viewed as providing some guarantee of the quality of the reference index. Having *academics* from all over Europe create lists of good scholarly journals in the humanities—rather than bureaucrats or metrics gurus—has given a new dimension to the project.

The active engagement of Expert Panels with the project has been crucial for the development of ERIH from its first conception in 2001. As Panels worked with the data supplied during the latter half of 2005, and reflected on the criteria established by the ERIH Steering Committee, they proposed some modifications to the original concept. ESF therefore created a forum for the Chairs of Panels to meet with each other, and also as a group with representatives of the ERIH Steering Committee in February 2006. Areas of concern for specific panels (disciplines) were addressed, for example the boundaries between the panels for archaeology and classical studies. The Panel Chair meeting also reflected more broadly upon the development of ERIH, and requested a wide consultation exercise for spring 2006, which was coordinated in the UK

by AHRC. Further discussions about the future of ERIH are under way regarding other issues as well.

Key developments are:

1) ESF acknowledges that ERIH cannot be a bibliometric tool. It is not possible or appropriate to use a categorised list of good research journals (compiled by criteria of quality alone) in a quantitative exercise such as bibliometrics. This would not be the best use of ERIH data.

2) Panel Chairs insist (and ESF agrees) that the use of the A/B/C categories of journals for the bibliometric assessment of departments or individuals for funding, promotion or employment, is inappropriate.

3) Panel Chairs emphasise (and ESF agrees) that the A/B/C categories of journals do not represent a hierarchy in quality. Some would have preferred to drop the A/B/C designations altogether, because of their misleading hierarchical implications—they were maintained in order to avoid confusion in the consultation process.

The guidelines for these categories describe different kinds of journals, serving different roles in academic exchange. All three categories comprise titles that are important for top-quality research. World-class scholars certainly publish research in all three types of journals, depending on the subject matter and the audience, as the author of 'Doing the sums …' rightly points out. The same holds true for the journals consulted and cited—inevitably scholars will draw on research published in all three categories. Moreover, the venue of publication offers no guarantee of the quality (positive or negative) of any individual piece of research. The author of 'Doing the sums …' uses simple statistics to provide a good illustration of this fact, broadly characteristic of humanities research. This is the very basis for the entire ERIH project.

4) The ERIH lists currently compiled represent only a first step. In order to be useful (and some possible uses are described below), they will need to be regularly reviewed and updated.

5) The predominance of monographs and edited collections of papers in humanities scholarship in relation to journals is recognized. Plans for the future development of ERIH take this into account.

Thanks to the strong engagement of academics in the development of ERIH, the project is also now evolving in new directions which will benefit humanities research in Europe in concrete and substantial ways:

1) Encouraging 'best practice' in the publication of journals.

Less an issue for UK and English-language journals, procedures such as peer review of articles are far from being standard in all traditions of humanities scholarship in Europe (and beyond). Indeed, in the context of Europe British scholarship overall stands out as very strong. Scholarly traditions are extremely varied across Europe as well as within different subject areas, and Expert Panels were well aware of this. However it was generally agreed that standards can be developed which would describe an 'international journal'. Such a journal should normally have:

- a genuine, varied and regular international cohort of contributors and readership;
- a consistently high-quality scholarly content;
- a broad consensus within the field concerning international status and visibility, insofar as possible;
- a quality control mechanism, normally through peer-review.

Panels struggled to capture the diversity of European scholarly traditions under this last point, by insisting that journals would depart from peer review only where there is another system ensuring quality control. With this phrasing, panels aimed at being sensitive to scholarly traditions in which peer-review is an unfamiliar procedure.

Panels hope that ERIH will encourage journals to declare their established transparent and accountable methods of quality control, and request for the next stages of the ERIH process an active engagement with editors and publishers to this effect.

Provided that resources are available, ESF should set up a mechanism to ensure that journals' quality control procedures are verified.

Beyond the four main characteristics, international journals will have some, though not necessarily all, of the following characteristics:

- *Active* international advisory board
- Openness to unsolicited contributions
- Highly discriminating and selective in the choice of articles published
- Publication on time and to an agreed schedule

It was suggested that the degree to which these

secondary characteristics applied could be added to the criteria distinguishing journals in categories 'A' and 'B'.

The expert panels believe, therefore, that one key aim of ERIH should therefore become the encouragement of high standards of professionalism in the publication of humanities research in European journals.

2) Disseminating European research in the Humanities worldwide

In the medium to long term, the identification of high-level venues of publication for scholarly journal articles in the humanities by ERIH has the potential for development as a powerful research tool, creating a showcase for European research in the humanities.

In the first instance, ERIH can be used as a platform for making information about journals and their contents more visible and easier to access worldwide, ideally through an online web portal. Options for funding are currently being discussed.

In a more visionary approach, ERIH could provide the basis for a European journal database, a kind of 'Euro-stor' providing accessibility and visibility for European multilingual scholarly research through:

- Development of user-friendly and reliable web portals
- Designing a long-term programme of digitalization
  - Making back issues of journals available
  - Translations of abstracts and ultimately of key articles
  - Development of key bibliographic tools such as EndNote codes
  - Encouraging European E-journal output
  - Developing tools and programmes for using ERIH in E-learning, VLEs and VREs

This would be a major step for the dissemination of research, notably for work which is not published in English and which is sometimes extremely difficult to access for those outside institutions with major research libraries. JSTOR already does an excellent job of covering a considerable amount of English-language scholarly output, but important research in other languages is poorly served by comparison. Given the degree to which language is at the heart of national and regional identities in Europe and

elsewhere, and underpins academic traditions in many fields, the availability of electronic dissemination is crucial.

There is much debate about the use possible of 'metrics' for the assessment of academic 'quality' in the UK and beyond. Academics and funders agree that use of any kind of metrics in a simplistic way is a grave concern, when the overall aim is to strengthen the quality and international visibility of humanities research all over Europe. We should be engaging in these discussions at a wider European level and ensure that large-scale projects, such as ERIH, can benefit from discussions here in the UK. More importantly, ERIH has the potential to evolve into something much more exciting and useful than what was originally envisaged, so let us make the most of it.

Addendum

The relationship between ERIH and the recently established HEFCE/AHRC Expert Group on Research Metrics.

- the HEFCE/AHRC Expert Group has the mandate to explore different forms of metrics considered as appropriate for the humanities, but does not base its discussions on the ERIH project; in fact its terms of reference explicitly rule out the use of ERIH
- the Expert Group, while committed to exploring possibilities for metrics, has certainly not ruled out peer review for outputs
- the Expert Group has engaged, on this and other aspects of its work, in a wide-ranging consultation exercise with major groups of academic stake-holders, including RAE Panel Chairs, AHRC Panellists and Panel Chairs, members of the AHRC Council, subject associations and HEIs
- ERIH is not considered, either by the HEFCE/AHRC Expert Group, or by ESF, its member organisations, its Steering Committee and other stakeholders, as being the work of 'those who propose metrical assessment to be the answer to the question of how to assess scholarly work'.

LIN FOXHALL
UNIVERSITY OF LEICESTER

CLASSICS AND METRICS-LED RESEARCH ASSESSMENT

Introduction

To evaluate the possible use of metrics for research assessment in Classics, we need to know *what* metrics, used *how*, and *for what purpose.* Appendix 2 of the Department for Education and Skills consultation document on the reform of higher education research assessment and funding[1] claims to describe a possible 'model' for funding in the Arts and Humanities, but in reality it presents a list of vaguely described metrics (e.g. 'bibliometrics'), all of which 'would deliver the quality assessment' in a way that is left wholly unspecified, and all or some of which 'could drive the funding formula', again in a way that is left wholly unspecified. The models which the document outlines for STEM subjects[2] are not nearly so vague. The restraint which this vagueness reflects is commendable; but it means that at present we do not have much basis for intelligent or informed discussion.

For *informed* discussion, we also need evidence, and that requires detailed research on the validity of different metrics, and the possible effects of their use for research assessment and funding, at the level of particular disciplines. So far as I am aware, no such research has been done. So it is hard to see how at present any answer to consultation question 4 ('What, in your view, would be an appropriate and workable basis for assessing and funding research in non-STEM subjects?') could be more than impressionistic guesswork. That seems like a dangerous basis on which to discuss important policy decisions.

It should be stressed that this *caveat* applies equally to rejections of metrics as a basis for research assessment and funding in the Arts and Humanities. No one would claim that the existing peer review system is perfect, and a metrics-based system that was genuinely 'appropriate and workable' would have great practical advantages. At present, we do not have solid grounds for asserting that such a system is impossible in principle. Nevertheless, it is reasonable to approach metrics with caution, since research could be seriously damaged by the adoption of a system that failed to assess research quality accurately, or produced perverse incentives. In what follows, I examine the prospects for the use of metrics in the assessment of research in Classics. Although I give most attention to bibliometrics, the discussion will lead to some broader conclusions about the use of other metrics.

Bibliometrics

Bibliometrics have been developed primarily in connection with STEM subjects. There has been some research in Social Sciences and Humanities, or Arts and Humanities, in general.[3] But it is not clear that these broad subject groupings are sufficiently uniform to be valid. Studies of individual disciplines are patchy. Since I could not find any bibliometric studies relating specifically to Classics, I decided to do a small-scale investigation of my own. This is an informal, exploratory study of a very limited data set, not conducted in a very rigorous (or statistically sophisticated) way; so I do not make any strong claims for its validity. But unless I have missed something in the existing literature, it is as good as we have got until someone does a larger and more rigorous study.

Procedure

My investigation had two goals. First, I wanted to test the commonplace observation that research in Classics may remain current for much longer than research in most scientific disciplines. Secondly, I wanted to get a sense of how citations of individual publications are distributed over their histories. I therefore proceeded in two stages:

---

[1] *Reform of higher education research assessment and funding* (www.dfes.gov.uk/consultations, 13 June 2006).

[2] I.e. the sciences, technology, engineering, mathematics and medical subjects. I take the STEM/non-STEM distinction over from the consultation document, but do not feel happy with it. On the one hand, it ignores the (very abstract) commonalities across all kinds of research; on the other (and more important) hand, it abstracts from the differences between individual disciplines which, as will become clear, are crucial to adequate assessment and funding models.

[3] There is a recent survey in A.J. Nederhof, 'Bibliometric monitoring of research performance in the Social Sciences and the Humanities: a review', *Scientometrics* 66 (2006), 81–100. Nederhof is alert to interdisciplinary variations, but, as we shall see, the conclusions which he reports, rather cautiously, for 'many' social science and humanities disciplines do not match the findings reported below for Classics.

28

First, I collected all secondary citations (excluding self-citations) from all articles published in *American Journal of Philology* in 2005, and analysed their chronological distribution.[4]

Secondly, I undertook a longitudinal study of the citation histories of a sample of publications. My initial sample consisted of the 14 classical items cited in *AJP* 2005 that had been published in 1986.[5] But this was a narrow and not necessarily representative sample (by definition, it included only items that had a citation history of at least 20 years). So I subsequently enlarged the sample to include all articles and short notes published in 1986 in *Harvard Studies in Classical Philology*, *Classical Antiquity* and *Classical Quarterly*.[6] In each case, I used the ISI Web of Knowledge Arts & Humanities Citation Index[7] to trace citation histories from 1986 to 2005. Self-citations, book reviews, and citations in bibliographical listings were excluded from the citation counts.

The scale of the resulting samples was as follows:

(i) The 20 articles in *AJP* 2005 cited 1116 items.[8]

(ii) The 14 items in the initial sample of items published in 1986 accumulated a total of 735 citations over 20 years.

(iii) Enlarging the sample added another 93 items, and another 305 citations. That is, the enlarged sample contained a total of 107 items, which had accumulated a total of 1040 citations over 20 years.

(iv) In the initial sample, 9 of the 14 items were books. Since the enlargement was based on journals, the number of books remained the same in the enlarged sample. There is no reason

to believe that either sample reflects the actual proportion of different categories of publication in Classics, which is (to me, at least) unknown.

We do not know how citation counts might be used to inform any future RAE. For the present exercise, I have assumed a continuation of the current 7-year survey period. This means that the citations counted would be citations within the 7-year survey period of items published within the 7-year survey period.[9]

Chronological depth

The distribution of citations in *AJP* 2005 by decade over a complete century was as follows:[10]

1996–2005	22.2%
1986–1995	28.4%
1976–1985	18.1%
1966–1975	11.7%
1956–1965	6.3%
1946–1955	3.6%
1936–1945	2.6%
1926–1935	1.5%
1916–1925	1.7%
1906–1915	1.2%

The remaining 2.7% of citations were spread thinly, all the way back to 1786.

Only 12.8% of the cited items cited in *AJP* 2005 were published in the 7-year period 1999–

---

[4] *AJP* was chosen, not only because it is one of the leading journals in the field, but more particularly because its house-style requires each article to have a bibliography, making it easy to extract the citations.

[5] The choice of 1986 was largely arbitrary: but 20 years is a round number, and I had also discovered that it was a landmark year in the distribution of *AJP* 2005 citations (see §9 below).

[6] *HSCP* and *CA* were the two significant Classics journals cited in the initial sample; *CQ* was chosen as another leading journal.

[7] Arguably, it would have been better to include the Science and Social Science Citation Indexes, so as to capture interdisciplinary impact. Previous experience of using the Citation Indexes suggests that this would give a modest boost to a limited number of items in the survey.

[8] Some items are cited more than once, but the distribution barely changes if one analyses the 1157 separate citations rather than the 1116 cited items.

[9] Something like this is more or less forced, since the further the period is extended backwards, the less valid the RAE becomes as an assessment of current research performance. It is the model assumed in M. Norris and C. Oppenheim, 'Citation counts and the Research Assessment Exercise V—Archaeology and the 2001 RAE', *Journal of Documentation* 59 (2003), 709–730. It is clear that Norris and Oppenheim are treating archaeology as a representative humanities discipline; the fact that the most cited articles in their survey were published in *American Journal of Human Genetics*, *Genetics*, and *Radiocarbon* should have given them pause. This is a striking example of the lack of understanding of disciplinary distinctiveness that gives bibliometrics a bad name.

[10] The lower rate of citation for the most recent decade reflects the time taken for new publications to be assimilated plus the time taken for new research to be published: unsurprisingly, very few (in fact, only 5) items published in 2004 and 2005 were cited in articles published in 2005. The analysis must be shifted two years back for the most recent decade to score more highly than the one that precedes it (29.5% to 24.0%).

2005.[11] To capture 50% of the cited items, one has to go back as far 1986. In other words, half of the citations are of items that are more than 20 years old.[12] The mean age of cited items was 27 years.

Importance of books

Around 53% of the 1116 items cited in *AJP* 2005 were books. In the initial sample of items published in 1986, 9 of the 14 items (64.3%) were books; these books accounted for 78.6% of the accumulated citations. Books were a much smaller proportion of the total in the enlarged sample; nevertheless, the 9 books (8.4% of the 107 items) accounted for 55.6% of the accumulated citations. Of the top 10 items in the enlarged sample, 7 were books (articles were ranked 4th, 8th and 10th), and all the books were placed in the first quartile. So even without a balanced sample (p. 29 above), this investigation confirms the importance of books as a medium of scholarly communication in Classics, and illustrates the relatively greater long-term impact of books as compared to articles.[13]

The importance of books is sometimes urged as an objection to using journal-based Citation Indices. That is a misunderstanding. Although the Arts & Humanities Citation Index only excerpts references *from* journal articles, it records references *to* books as well as articles. There would be a distortion if citation patterns in books and articles differ in a systematic way. It is entirely possible that there is no systematic difference; but that is, at present, an untested assumption.

Citation histories

Both the initial and the enlarged samples showed a peak in the number of new citations in years 8–10; i.e. the peak in citation rate occurs beyond the horizon of a 7-year RAE survey period. But closer analysis shows that the 8–10 year peak is a side-effect of the concentration of citations on books. The books peaked in years 8–10; the articles peaked in years 3–5. However, in each case the peaks were very modest. Books, articles and the sample as a whole all continued to accumulate citations at a remarkably steady rate over the 20 years. (In each case, r-squared for the linear trend line exceeds 0.99.) Consequently, even articles had accumulated only 38.5% of their 20-year citation total by the end of year 7 (books 23.0%, total sample 29.9%).

Although the citation rate was steady for the whole sample, the rate at which individual items accumulated citations over the 20-year period was highly variable. This means that extrapolation from the 7-year count is not a reliable predictor of an item's long-term impact. For example, of the five articles that had accumulated 8 citations by year 20, three had 3 citations after year 7; but one had 5, and one had none (it was still two years short of its first citation).[14] How much this unreliability would matter in practice depends on how the data might be used to drive quality ratings and funding allocations in a real RAE, which is completely unknown. For illustrative purposes, I have assumed that the 20-year citation count provides a reasonable indication of each item's long-term impact, and have used the mean and median number of citations as performance benchmarks. That is, I have posed the question: how often is performance relative to the mean after 7 years a reliable predictor of performance relative to the mean after 20 years?

(i) The mean number of citations per item after 20 years was 9.7; 17 items achieved this benchmark. Performance relative to the mean after 7 years gave a 15.9% error rate: 1 false negative (out of 17 = 5.9%), and 16 false positives (out of 90 = 17.8%)).

(ii) The median number of citations per item after 20 years was 3; 56 items achieved this benchmark. Performance relative to the median after 7 years gave a 17.8% error rate: 5 false negatives (out of 56 = 8.9%), 14 false positives (out of 51 = 27.5%).

---

[11] This is equal to the citations from the period 1900–1959; a further 2.5% of citations were from before 1900.

[12] Hence Classics lies outside the range of disciplines covered by Nederhof's observation that 'reference lists in social sciences and humanities articles more often contain publications older than 5, 10, *or even 15 years* than in the sciences' (2006: 86, emphasis added).

[13] More cautiously: of books that have a citation history of at least 20 years as compared to articles in general.

[14] More precisely: it was two years short of the first citation recorded in the Arts & Humanities Citation Index. Since the Index is very limited in its coverage of Classics journals, and it does not record citations in books at all, the *absolute* number of citations per item is certainly understated here. (So the fact that 16 articles had recorded 1 citation, and 22 had recorded 0 citations, after 20 years, should be treated with caution.) Here, however, the point is the unpredictable variability in the *relative* number of citations per item over time.

In view of the evidence (p. 30 above) that citation of articles peaks earlier, I tested these benchmarks on the 98 articles separately. The test showed that the 7-year citation count is even less reliable for articles (despite being less discriminating):

(i) The mean number of citations per article after 20 years was 4.7; 31 articles achieved this benchmark. Performance relative to the mean after 7 years gave a 16.3% error rate: 4 false negatives (out of 31 = 12.9%), 12 false positives (out of 67 = 17.9%).

(ii) The median number of citations per article after 20 years was 2; 60 articles achieved this benchmark. Performance relative to the median after 7 years gave a 15.3% error rate: 10 false negatives (out of 60 = 16.7%), 5 false positives (out of 38 = 13.2%).

It should be noted that the figures given in the previous two paragraphs understate the scale of the problem. Most of the items published in a 7-year RAE survey period will have been published in years 2 to 7, and so will not have accumulated a full 7 years of citations. Their interim citation counts would therefore be even less reliable predictors of long-term performance.

The conclusion to be drawn from these figures is that, as a consequence of the relatively slow publication cycle in research in Classics and the very long shelf-life of published research, long-term impact is not accurately reflected in, and cannot be reliably extrapolated from, citation counts over a period short enough to be appropriate for an assessment exercise.[15]

Journal Impact Factors

If citation counts within a 7-year survey period are not a reliable direct indicator of long-term impact, there may be a temptation to use a proxy to predict long-term impact. The most obvious proxy is journal Impact Factors. The rationale would be that journals' peer review mechanisms will already have sorted articles into a qualitative hierarchy that roughly corresponds to the qualitative hierarchy of the journals themselves, which is reflected in their Impact Factors. In fact, different articles in the same issue of a journal vary widely in their frequency of citation, so this use of Impact Factors is of questionable validity in general. But there are also discipline-specific issues.

The Institute for Scientific Information calculates journal Impact Factors on the basis of the number of times articles published over a two year period are cited in the third year (e.g., the number of times that articles published in 2003–2004 are cited in 2005). Obviously, this is a hopelessly short time-scale for Classics, and a completely different approach to the calculation of impact factors would be needed.

Even if a different basis of calculation were adopted, there would be problems with applying this approach in Classics, given the pattern of scholarly communication in the discipline:

(i) Impact Factors will be poor discriminators in disciplines in which the journal hierarchy is relatively flat. I have not done any research to test my impression that this is the case in Classics. But since RAE panels in Classics have consistently refused to use location of publication as a criterion of quality, the hypothesis evidently conforms to the perceptions of experienced researchers in the field.

(ii) Journal Impact Factors are unlikely to provide a useful metric in disciplines in which books are an important medium of scholarly communication. We have already seen that this is the case in Classics (p. 30 above).[16]

(iii) The calculation of journal Impact Factors will be distorted by differences between journals in the extent to which they carry items that are likely to be cited less often, such as short notes on textual problems.[17]

---

[15] Again, Nederhof's conclusions seem to understate the position for Classics: 'Short-term citation impact tends to give a reasonable indication of medium or even long-term impact in many fields … In some fields, or parts of fields, however, longer 'mid-term' citation windows may be needed, ranging for example, from 1–4 years to 1–6 years instead of 1–3 years … For non-journal material, such as monographs and chapters, citation peaks tend to occur relatively late. Both in science disciplines and in social sciences and humanities, a five-year citation window tends to be a minimum requirement for meaningful analysis' (2006: 93).

[16] Note that the problem is not whether books will be captured in citation counts (see p. 30 above). The problem is that, if the unreliability of citation counts prompts us to move to journal Impact Factors, this cannot be applied to books. That would be a major problem in Classics, though entirely irrelevant in (e.g.) particle physics.

[17] Citation counts will fail to distinguish an obviously flawed conjecture that all subsequent researchers ignore from an obviously successful conjecture that is at once taken up in the text and apparatus of a new standard edition. The long-term influence of a publication may be considerable but, because it is largely mediated, invisible to

It is possible that these problems would be negligible at some suitably large level of aggregation. The problem is, again, that we have no idea what metrics would be used, for what, or how. For example, it is unlikely that unreliability at the level of individual items would be significant if bibliometrics are used to benchmark Classics in the UK against global competition.[18] On the other hand, this unreliability will be much more problematic if bibliometrics are used to rate (and allocate funding to) individual departments.[19]

Tests with arbitrary subsets of my data on the scale of a small-to-medium sized department's RAE submission suggest that this is too low a level of aggregation to avoid substantial variations in the extent to which 7-year citation counts accurately reflect long-term impact. If that is the case, bibliometrics would be inappropriate as determinants of quality ratings and (especially) funding allocations at the level of individual departments.

At the beginning I emphasised the differences between disciplines within the Arts and Humanities grouping, and the consequent importance of *discipline specific* research and *discipline specific* metrics. Realistically, we cannot expect RAE metrics to be so finely tuned. This, too, has implications for the appropriate level of aggregation. To the extent that a metrics-based system is not discipline-specific, it should only be used to generate funding allocations at a level of aggregation sufficiently high to drown out the distortions they will produce at the level of individual disciplines.

Possible side-effects

It should also be clear from what has been said that the use of bibliometrics as a basis for funding allocations is likely to reward patterns of publication behaviour different from those currently practised within the discipline—the more so, the less finely tuned those metrics are to specific disciplines. This is fraught with danger. There is a presumption that different communities of researchers have evolved patterns of publication behaviour which experience has proved to be most appropriate to their respective disciplines. Classicists and particle physicists have evolved very different patterns of scholarly communication in response to the very different communication needs of their discipline. Changes in publication behaviour that are driven, not by intrinsic academic factors, but by the extrinsic factor of a funding body's choice of metric, are unlikely to be academically beneficial.[20]

As an example, consider the possible consequences of a funding regime that rewards publication in journals with the highest impact factors. That is likely to increase the number of submissions to those journals. That, in turn, will increase their rejection rate. As a consequence, there will be an increase in the overall number of submissions in the system, as articles are resubmitted further down the hierarchy. That will require more editorial and peer review effort; and it will change the nature of the reviewer's task (the question will no longer be, 'is this good work, worth publishing?', but 'is this work good enough to be published in *this* journal?'). Hence the total peer review workload would substantially increase. That cost may be worth bearing in disciplines in which pressures such as publication volume and rapid turnover make a steeply hierarchised prior grading of outputs essential to efficient communication (although there seems to be a growing feeling in at least some scientific subjects that this is placing the peer review system under unsustainable strain). Disciplines in which those pressures are weak have no intrinsic reason to bear the additional cost.[21]

---

bibliometric surveys. This is a further general problem with citation counts in Classics.

[18] But the current Citation Index would be hopeless for this purpose. The bias of journal coverage to US/UK publications means that the citation habits of non-Anglophone researchers (which are likely to be different from those of Anglophones) will be under-represented.

[19] See §4.4 of the consultation document. But this confusingly runs together the separate issues of level of aggregation and the inadequate coverage of Arts and Humanities journals in the Citation Indexes: there is no reason to think that addressing the latter will have any bearing on the former.

[20] See §3.6, §5.21 of the consultation document for awareness of the potential undesirable behavioural consequences of an assessment/funding allocation model.

[21] Furthermore, the trend would be self-reinforcing. The more selective the top-rated journals are, the higher their impact factors, and the greater the incentive to publish in them. This will, in turn, have an impact on readers' habits: the more steeply hierarchised the journals, the less likely it is that research published in less highly rated journals will be noticed—irrespective of the actual quality of the research. Unless one is very confident in the ability of reviewers to predict the importance of submissions, this must seem an undesirable consequence.

This point is parallel to the problem that (as many have pointed out) would follow from the use of research grant income as a basis for allocating QR funding in Arts and Humanities. If research grants are needed, not simply to fund specific projects, but also to secure the research base, competition for grants will intensify. The resulting increase in the number of applications (a far higher proportion of which will inevitably be unsuccessful) will be enormously wasteful of the time and resources both of applicants and of grant-awarding bodies. Consider, e.g., the AHRC Research Leave scheme, which involves (i) the applicant, (ii) a nominated referee, (iii) two peer review assessors, (iv) the applicant again (responding to the assessors' comments), (v) the AHRC subject panel, and at various stages (vi) University and (vii) AHRC administrators. I think this is a good system (going through it substantially improved my thinking about my own current research project): but would it be sustainable in the face of a substantial increase in applications?

Quality-metric correlations

Advocates of the use of bibliometrics in research assessment often point to the significant correlation that has been established in a number of disciplines between the results of bibliometric analyses and the quality ratings produced by peer reviewers in past RAEs.[22] Let us assume, for the sake of the argument, that a similar correlation would be discovered if and when the analysis is extended to Classics. Clearly, the existence of a correlation is important: if a metric *failed* to correlate with the qualitative judgements of experienced experts in the field, that would tend to discredit the metric. But it is important to understand that demonstrating a correlation is not *sufficient* to validate a metric.

First, a fundamental point about statistical correlations: they do not on their own disclose the causal relationships between the variables. Therefore, a correlation cannot predict the consequences of a change in the structure of those relationships. So a past correlation between quality ratings and citation counts cannot predict what will happen when a funding incentive that has hitherto been attached to quality ratings is reassigned to the citation counts. To predict that, one has to understand the causal relationships underlying the correlation. After all, the only thing that is likely to achieve a high quality rating is high quality research; but citation counts may be influenced by things other than research quality.

In Classics, it is very obvious that the quality of the research is not the only factor that influences the number of citations an item receives. Researchers only cite good research if it is relevant to what they are doing. So one would expect research on intensively discussed topics to be more frequently cited than work of equal quality on a less intensively discussed topic. For example, the on-line *L'Année philologique* reveals that publications on Homer outnumbered publications on Heliodorus in the decade since 1995 by more than 30 to 1. Given such a disparity in the number of *opportunities* to be cited, it is inevitable that research on Homer will receive more *actual* citations than research on Heliodorus.

To test this prediction, I applied the methodology described above to my own work. Comparing the number of times my 1985–1991 publications were cited in the period 1985–1991 with the number of times my 1999–2005 publications were cited in 1999–2005, I found (not to my surprise) that there were almost six times as many citations in the former period. This does not reflect a sixfold decline in the quality of my research over the course of my career! But in 1985–1991 I was publishing on tragedy, Aristophanes and Pindar; in 1999–2005 I was publishing on the Greek rhetoricians of late antiquity. This shift of attention from texts that are intensively discussed to texts that are largely neglected has had a predictably adverse effect on my citation counts.[23]

---

[22] E.g. Norris and Oppenheim (2003).

[23] Experts in bibliometrics recognise the point at stake here. E.g. Nederhof: 'Citation impact indicators (such as the number of citations per publication) need to be compared with the citation impact of ... other publications in the same subfield' (2006: 93); Anthony F.J. van Raan, 'The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments', *Technikfolgenabschätzung—Theorie und Praxis* 12 (2003), 20–29: 'research fields should never be compared on the basis of absolute numbers of citations. Field-dependent normalization is absolutely necessary' (p. 25). However, I am not sure that they realise how fine-grained the discrimination of research fields would need to be; and I am quite certain that this level of fine-tuning would not be attempted in RAE.

This brings us back to the question of side-effects (cf. p. 32). If funding allocations become linked to citation-based bibliometrics, it would be irresponsible of me to damage my department's future income by working on a topic that would inevitably attract relatively few citations. Not everyone, perhaps, will think that a continuing neglect of later Greek rhetorical theory would be catastrophic for the discipline, but the point can be generalised. A funding structure that rewards researchers in proportion to the number of other researchers working on the same topic would encourage an ever increasing narrowing of the focus of research towards the most intensively (excessively?) discussed topics. That certainly would not be beneficial to the discipline as a whole.

The important conclusion to draw from this is that we ought not to be too impressed by the discovery of a correlation between quality ratings and metrics. For the dismal scenario of research being ever more narrowly focused would not upset that correlation. In fact, it would *improve* the correlation. If the system creates a deterrent against research on topics that are not intensively discussed, the phenomenon of high-quality low-citation research will disappear; all the remaining high-quality research will also be high-citation. Under such circumstances, a correlation of quality and metric would in reality be a sign of the discipline's poor health!

Other metrics

The HEFCE/AHRC joint working group on the use of metrics in the assessment of research quality and the funding of research shares my scepticism about bibliometrics: 'In response to questioning, the Chairman clarified that the expert group would not be recommending the use of bibliometrics or citations in its final report, although it understood that this was a rapidly developing science. It would recommend that a sceptical eye be kept on developments.'[24] But I am not much cheered by this.

I have focused on bibliometrics here because they provide the *most plausible* metric for estimating research quality. High-quality research is likely, other things being equal, to be cited more frequently than low-quality research. So there is no doubt that a genuine causal relationship exists between research quality and citation counts. Unfortunately, the analysis presented here suggests that this relationship does not yield a viable metric, for two main reasons. First, the relationship unfolds over too long a time-scale to be reliably captured in a plausible RAE window. Secondly, the rate of citation is too sensitive to variations in the level of research activity for citation counts to provide an index of research quality that is consistent across different topics within the discipline. Thus citations provide a metric that is plausible in principle, but unusable in practice. By contrast, other metrics do not look plausible even in principle.

Again, discussion is hampered by the absence of any clarity about which metrics might be used, and how. But it should be fairly obvious that the quantification of 'esteem indicators' on anything other than an arbitrary basis presents a formidable methodological challenge. And how will the result be meaningfully commensurated with (for example) research income or post-graduate student numbers? Do we even know whether these things provide valid indicators of research quality in Classics? I have argued that correlation is not sufficient to establish a valid metric, but it is certainly a minimum requirement. A metric that does not correlate with expert qualitative judgements will not have credibility. But analysis of the data for the 2001 RAE suggests that ratings were very weakly correlated indeed with research income, PGR numbers, PGR studentships, and PGR awards (either total or per staff FTE).

Research income may be worth a closer look, since the logic of using 'input metrics' is particularly opaque. Competition for research grants is not a mechanism for income generation. The competition distributes, but does not increase, the total funding available in the system. So the competition for grants does not in itself bring any net benefit to the discipline as a whole (indeed, it incurs significant costs: p. 32). The benefit to the discipline comes through the research which the grants enable. Consequently, if research income is used well, the beneficial effect will show up in (and be rewarded through) output metrics; if it is not used well, it ought not to attract further reward. Furthermore, while it is important there should be a mechanism for enabling inherently expensive research by meeting its costs through direct project grants, it is hard to see why those grants should have the

---

[24] HEFCE summary of the *Consultation seminar to discuss metrics in the arts and humanities* (www.hefce.ac.uk/research/assessment/reform/consultation.pdf, 25 September 2006).

additional indirect effect of increasing QR allocations. The production of high-quality research is not inherently more meritorious for being expensive.

RAE panels in Classics, past and present, have held that research assessment and funding should be driven predominantly by research output; in RAE 2008, output is weighted at 80%. A system that gave output a subordinate role (for example, 30%)[25] would therefore lack credibility. It would diverge sharply from the judgement of experts in the discipline as to the appropriate criteria of research excellence. It would sharply increase the rewards for things that have not been found to correlate with expert assessments of research quality in the past. And we have already seen how the use of research income, in particular, as a basis for allocating QR funding might have undesirable systemic consequences (p. 32 above).

The joint working group has 'looked at the entire range of metrics and data-sets available, with a view to determining which elements could be deployed within an assessment system for the arts and humanities in the short to medium term'.[26] This is not reassuring: 'looking at' does not sound much like rigorously conducted research. The group will recommend that 'whatever set of metrics is agreed, the outcomes produced by a metrics-led model should be compared with the results of the RAE 2001 and 2008 to help assess their validity'.[27] Would it not be better to assess their validity *before* they are agreed? Indeed, would it not be better to demonstrate that some valid set of metrics actually exists before making a commitment to the model being metrics-led at all?

Concluding comments

A few years ago the AHRC expressed robust scepticism about metrics: 'Plausible metrics-based systems of assessment for arts and humanities research would be impossible to construct ... Metrics cannot be the basis of a quality assessment that is to carry any credibility

in the arts and humanities ... There are fundamental problems with using bibliometrics and external grant income as surrogates for quality.'[28] For reasons explained earlier (p. 28), this strikes me as over-dogmatic. Nevertheless, I am not aware that any *academic* grounds for thinking that metrics carry credibility have come to light in the intervening years. That may explain the retreat to a much weaker criterion in the remit of the joint working group: 'to advise ... on what metrics-based approaches are possible now for these purposes, or might become possible in the next few years.'[29] But no one ever claimed that using metrics was not *possible*!

The CUCD's summary of the 25 September consultation seminar reports the Chair of the joint working group, Michael Worton, as claiming that 'we cannot and must not ... make a case for A&H being distinctly different from the STEM areas'.[30] The claim is puzzling. Why cannot and must we not make a case for something that is clearly and consistently recognised in the government's own consultation document?[31] If the government spontaneously acknowledges what most researchers in the Arts and Humanities maintain, that looks like an opportunity too valuable to be squandered. The claim was advanced at the consultation seminar that 'there was a danger that making the arts and humanities seem exceptional in this respect would undermine parity of esteem with science, technology, engineering and mathematics (STEM) disciplines, and could ultimately lead to funds being diverted'.[32] Possibly—though the consultation document gives no support to the claim. But there is certainly also a danger that applying inappropriate metrics will inflict systemic damage on the conduct of Arts and Humanities research, which one might reasonably feel to be more pressing. To say that 'there is a large area of common concern across

---

[25] A model proposed ('to provoke debate') by the joint working group assigns a 30% weighting to outputs; research income carries another 30%, with a mix of other factors accounting for the remaining 40%. See A. Lipsett, 'Arts metrics plan revealed', *THES* 8 September 2006, p.4.

[26] Tony McEnery (AHRC), as reported in the HEFCE summary.

[27] Michael Worton (Chair of the working group), *ibid.*

[28] AHRC, *Response to the Funding Councils Review of Research Assessment* (http://www.ahrc.ac.uk/images/4_92095.pdf, November 2002).

[29] HEFCE/AHRC joint press release, (www.hefce.ac.uk/news/hefce/2006/metrics.htm, 4 July 2006).

[30] In the HEFCE summary, this is much more cautiously expressed: 'it would be *preferable* if all research activity were treated as a single spectrum for assessment purposes' (added emphasis).

[31] §1.9, §3.2, §5.3, §5.14, §5.19.

[32] Michael Worton, as reported in the HEFCE summary.

the whole academic sector'[33] is self-evident, but not to the point: what needs to be shown is that there are no relevant differences—which seems to me patently false.

We do not currently have evidence to prove that it is impossible in principle to devise a metrics-based system for research assessment in the Arts and Humanities that would be no less robust than the current (and admittedly imperfect) assessment regime. But since no model has yet been proposed that looks remotely plausible, it is reasonable to remain sceptical until the validity of some model has been established.[34] And that, as I said initially, requires research. The one thing that *is* certain is that discussing the possibility of metrics-led research assessment is pointless and dangerous if the discussion is not informed by rigorous evidence-based research at a discipline-specific level, since the resulting *uninformed* system is likely to have a distorting influence on researchers' behaviour. The AHRC, which should be taking the lead in filling the evidential gap, has little motive to do so, given its willingness to maintain (for reasons that are apparently political rather than academic) that there are no relevant differences between STEM and non-STEM, or within non-STEM, disciplines, and its willingness to embrace the principle of a metrics-led model *before* the validity of any metrics-based model has been established. At present, the prospects for an informed outcome to this debate do not seem good.

Addendum

This paper was written before the publication of the Expert Group's report, *Use of research metrics in the arts and humanities.*[35] I append some brief commentary on the Group's findings.

[33] Michael Worton, as reported in the CUCD summary. Tony McEnery (AHRC) is reported as saying that 'the arts and humanities were distinctive in their concerns and approaches, but not so different from other disciplines that a separate system of assessment is warranted.' But what is the evidence? Where is the research on which this claim is based?

[34] One might still feel that a system which depends on the expert judgement of people who know the discipline well from the inside is less likely to create perverse incentives than one which is based on a quantitative formula.

[35] www.hefce.ac.uk/research/assessment/reform/metrics.pdf

The Group maintains that 'there is no *fundamental* difference in the nature of the research enterprise in the science, technology, engineering and mathematics (STEM) disciplines on the one hand, and the arts and humanities on the other' (Principle (iii), added emphasis). That is not contentious: there is clearly *some* level of generality at which it is possible to describe all research activities with a single formula (e.g. they all involve trying to find things out). The only possibly contentious question is whether this very abstract fact is relevant to the design of appropriate discipline-specific assessment regimes. But since this Principle does not have any discernible influence on the content of the rest of the report, that question can be set aside.

The body of the report begins in a way that suggests that metrics will play a major role in the proposed assessment regime: 'the assessment framework should incorporate a range of interrelated metrics that, in combination, will be an efficient and robust component of a system to assess research quality ... the assessment framework should be delivered through significantly lighter-touch peer review and greater use of metrical data' (Paragraph 11).

However, the conclusion is a little more restrained: 'at this point, metrics alone will not allow the overall performance and quality of research to be assessed at individual or departmental level. Nevertheless, metrics have an important role to play in research quality assessment' (Paragraphs 30–31).

But even this modest claim is unsubstantiated by the report's contents. Of the six elements listed in Paragraph 14, only three are intrinsically quantitative: (ii) Spend on research infrastructure and other funding of the research environment (although the inclusion of narrative self-assessments compromises the quantitative nature of this element); (iv) PhD completions per research active member of staff; and (v) Peer-reviewed external research income. However, the report recognises that neither PhD completions nor external research income can provide a useable metric without 'normalisation'. This recognition is welcome. But is any non-arbitrary method of normalisation available? The report does not adduce any evidence that there is. Although the Group acknowledges that 'outcomes will vary depending on what is counted and how it is counted', there is no evidence in the report that serious thought has been given to the difficult

questions of quantification and commensuration that are fundamental to establishing a 'robust' metric (above p. 34). Typically, the report falls back on modal auxiliaries when such questions arise, making vague gestures towards what 'would' be needed or 'should' be done. In sum, the report identifies a limited range of potentially relevant quantitative data, but gives no indication of how the data might be converted into useable metrics.

The relevance of some of the quantitative data is, in any case, questionable. In particular, it is difficult to see why spending and income figures should be seen as relevant indicators. These figures measure the consumption of resource. But consumption is not valuable in itself: it is valuable to the extent that the resource is used effectively. That cannot be inferred from quantitative data on spending and income. There is a risk that treating these as assessment criteria may reward (and therefore incentivise) the inefficient consumption of resource (above pp. 34–5). The report fails to explain how this risk is to be avoided.

As for the non-quantitative elements, the report does not go beyond vague hand-waving with regard to (iii) Wider social, cultural and economic significance of research process ('each broad disciplinary area *would* develop a suite of indicators which *could* provide robust evidence') and (vi) Esteem indicators ('this element does not lend itself easily to quantification'—indeed not: 3 invitations to contribute to edited volumes + 1 FBA = ?).

The Group's only substantive proposals relate to (i) Research outputs. First, the significance research outputs will be reduced: a uniform weighting, for which the Group suggests 50%, will be imposed across all disciplines (Paragraph 15). The suggested weighting seems entirely arbitrary, and is much lower than that which review panels, given the freedom to choose, have actually chosen (above p. 35). Second, the basis of the assessment will be made narrower by a move to sampling of submitted outputs (Paragraph 19). There is, however, no explanation of how the sample would be determined, and how the fairness of the sample would be ensured. The possibility that 'it would be possible to commission more than one peer review for a sampled output' (Paragraph 21) is not particularly attractive if this simply means that a researcher's best work might be ignored by more than one reviewer.

'Whatever the precise nature of the metrics chosen and their individual weightings, it is essential that a trial of their use be compared with the results of RAE 2008, so that the efficacy of a metrics-informed process can be evaluated' (Paragraph 12). The implicit assumption that the metrics can be evaluated by comparing their results with the results of non-metric assessments is dangerously flawed: one must consider, not only the correlations, but also the causal relationships underlying the correlations; otherwise, it is impossible to predict the effects that reassigning the incentive will have on researchers' behaviour (above pp. 33–4). And I am still left wondering: would it not be better to evaluate the metrics *before* making a choice (above p. 35)?

MALCOLM HEATH
UNIVERSITY OF LEEDS

WHAT CAN WE LEARN FROM THE USA?

Is this a question that is raised by scholars in other disciplines or by scholars beyond the UK?[1] Does it signify openness on the part of the UK classics community to the outside world and a willingness to learn other approaches and methods? Or does it point to an insecurity and a feeling of inferiority and to a sense that departments in the UK are being left behind by more 'advanced' developments overseas (especially in the USA)? I am tempted to quote a favourite philosopher, Jacques Derrida, and say that pedagogy and scholarship should have no passports and that the best forms of pedagogy and scholarship find a home and an audience everywhere.

But perhaps it is true that there is something distinctive about the pursuit of Greek and Roman antiquity in the UK and that there is a distinctively British idiom in which the areas that commonly fall under the rubric of 'Classics' have flourished. The history of scholarship suggests that this might be the case. And the barriers that separate modern nations and peoples from each other, even in the age of globalization, have made it harder for outside influences to come in and have promoted a British phenomenon. Moreover, the difference between British classics and other forms of the discipline may be quantitative as well as qualitative. Let us not forget that the resources devoted to classics and the training offered in several schools and universities in the UK are still impressive, even now in 2006, and difficult to match in most other countries. In fact, the UK, along with a handful of nations, offers a distinctly privileged space for the study of Greek and Roman antiquity, though this may not always be evident to those who are involved with it on a daily basis.

Here, then, are a few thoughts in answer to the question posed above. My answers are inevitably subjective and, undoubtedly, they are based on generalizations that can be disproved without much effort. I was educated in India and the USA, and I have taught in the USA and the UK. Since I have taught in the UK for only a year, and since my prior time in the UK has been limited to a visiting appointment of a few

months' duration, my remarks below may well suffer from a lack of familiarity with the domestic scene. Nevertheless, I offer these comments in the hope that, for all their shortcomings, they may stimulate thought and discussion. US classics departments could learn much from UK departments too, but I have for the most part avoided commenting on the matter and stuck to my brief in this essay.

Similarities

It might be proper to begin by acknowledging that there are, in fact, many similarities between classics departments in the UK and the USA. In both countries, departments offer classes that require no knowledge of Greek and Latin as well as classes that do require such knowledge, and in both countries, several departments offer classics (or 'classical civilization', or some variant thereof) degrees to students without requiring them to study any Greek or Latin. There will be classes in literature and history; there will be classes on specific authors; and there will be classes on mythology and art. We could consider numerous other administrative, intellectual, and pedagogic similarities, but let us move on to the differences and what we might learn from them.

What is Classics?

In many venues in the USA, classics departments are small, and colleagues often complain of marginalization or voice the fear that their departments may be shut down by the university. The public sphere does not automatically or necessarily grant classics the honoured status that it still enjoys in Western Europe. As a result, American colleagues often need to make a case (to their students, to their colleagues, to the university administration, and to the public at large) for the study of Greek and Roman antiquity and to provide an intellectual justification for their discipline. Having to explain the importance of one's work is sometimes helpful to one's own intellectual formation: it's a way to avoid complacency and to sharpen one's perspective on the subject. Of course, for classical scholars in the UK there are benefits that follow from the fact that Greece and Rome occupy a special place in European culture, and I do not mean to minimize these to the extent that they lead to the promotion of a lively and critical approach to

---

[1] This question was posed by CUCD Bulletin.

antiquity. For instance, the relatively large number of British theatre and television productions (even when these are of questionable merit) based on Greek and Roman antiquity cause people to engage with the past and to reflect on it further. These productions also often feature the views of classical scholars and writers and thereby give them a platform from which they can facilitate intelligent discussion.

Inter-disciplinary contact

My impression is that there are more contacts between departments in general and that there is a greater level of inter-departmental contact in the USA. Such a situation allows students and teachers in classics departments to benefit intellectually from other fields and disciplines, and it may be a phenomenon that UK departments wish to study. Interdisciplinary work is actively encouraged and rewarded in the USA, and it is easier for colleagues to reach across disciplines in their teaching and research. Classics classes often include large numbers of students who are non-majors. Perhaps those in the UK who now teach in 'Schools' (i.e. groups of departments) are also experiencing a healthy level of contact with colleagues in other areas. In general, I would like to make a plea for a greater degree of openness to other disciplines, approaches, methodologies. In that respect, the spread of reception studies in UK departments is a notable and welcome development.

Placement & Jobs

The search for academic jobs is more centralized in the USA than in the UK. The preliminary interviews routinely take place at one venue, the annual meeting of the American Philological Association, which runs a placement service. The meeting is held in early January; by late February, a candidate generally learns how successful s/he has been in the search and then decides on a course of action based on the offers in hand. I believe that UK departments and the Classical Association should explore the possibility of holding a placement service along the lines of the APA. The UK system is tilted in favour of the institution, while the US one in favour of the candidate. In the UK system, a job-seeker may secure a position in January and then learn about a more suitable position two months later. Surely, it is better for a candidate to have a full sense of the jobs that are going to be available in an academic year? And surely it is better for a candidate to apply for all relevant jobs, secure all relevant interviews, and then accept the best available option at the end of the process? I recognize that the universities have different financial constraints on them here, but pressure can be applied, at the local and national levels, to university administrators, vice-chancellors, and deans to follow a centralized system.

American departments also do not rely on two elite institutions for the supply of staff and hire from a wider group of universities. British classics departments hire their staff very heavily and disproportionately from Oxford and Cambridge. To some extent, such a domination of appointments is appropriate since the two universities produce the largest number of classics graduates and graduate students. But there are numerous departments outside of Oxford and Cambridge that educate brilliant students, who do not seem to have much success in finding jobs. Admittedly, this is an old problem—which also has an American counterpart—but it is still in need of a solution.

Teaching

The number of weeks spent in the classroom is greater in the USA than in the UK (30 vs. 22–24), and the period for examinations is much shorter. Examinations are held at the end of each term rather than at the end of the year. Students also need to complete many more assignments (including papers and midterm examinations) during the term, and only a small percentage (say, 30%) of their overall grade/mark for the class depends on the final examination. I believe that reducing the heavy emphasis given to final examinations here is a step in the right direction, and support the many departments who already follow the practice; it encourages students to participate more regularly in class and gives them an incentive to stay current with the syllabus and with reading assignments.

The division of labour could be more uniform in the UK. In most American research universities, and in many smaller American colleges, the teaching load is generally four or five courses a year, depending on whether the institution's year is divided into semesters or quarters. The teaching load is higher at less wealthy institutions. In the UK, my understanding is that teaching loads vary considerably from institution to institution and even within institutions. Departments in the UK

may wish to establish regularized and reasonable teaching loads for their staff.

Research universities in the USA often hire instructors from the ranks of their graduate students as Teaching Assistants, who help with the marking in large classes and run weekly discussion sections. In some cases, and especially with language classes, graduate students also serve as regular instructors. The use of graduate students in such a capacity gives students money and valuable teaching experience and gives the established staff a more practicable teaching load. I believe that such a practice already exists at some British universities; I would endorse it and add that the process could be more regularized and transparent.

Lastly, there is in the UK a large degree of oversight and supervision in relation to teaching and marking; this is accompanied by substantial amounts of paperwork. These can be brought to more manageable levels. There is seldom any double-marking or blind double-marking in American universities and no comparable use of external examiners. In these matters, I prefer the American system to the British, though I can see that the latter has its advantages.

Administration

I will now make a couple of general points that relate more to British universities than to classics departments but I think that the issues raised might partly be addressed at the departmental level as well.

If you think that you are forced to do a higher share of mundane administrative tasks in the UK than your counterpart in the USA, you are right! The administrative load assumed by academics in Britain is simply too high, and there is no way to deny this or to pretend otherwise. Universities in the US are more centralized and much of the administrative work there is carried out by administrators rather than by professors. Even departmental offices in the USA undertake far fewer administrative and bureaucratic tasks than their counterparts in the UK. It is the registrar's office, typically, that worries about class enrolments and waiting lists. Students are advised not just by counsellors or tutors but also by a centralized office the main purpose of which is to give guidance to students about academic programmes; this practice takes some

of the pressure off advisers within departments. However, it is not just the advising of students that is handled by professionals, but a whole range of administrative activities. The paperwork is significantly less in US universities.

Universities in the USA are subject neither to a Research Assessment Exercise (RAE) nor to a periodic review of teaching and learning activities, though there are major departmental reviews every 7 to 10 years at large institutions. The RAE and the periodic review place a significant burden on a department's resources, and, as every British academic says, should be implemented in ways that are far less disruptive. In the USA, however, there is a system for tenure, which brings its own share of problems, and I believe that UK institutions are better off without the tenure system!

Salaries

The UK has a national pay scale and most universities follow it. There is less opportunity here than in the USA for the university to offer different salaries to its staff. There is, if you will, a kind of national parity when it comes to salaries in the higher education sector. In the USA, both private and public universities follow no explicit national grid in determining the salaries they offer their staff, and as a result there is a huge degree of variance in faculty salaries, both within an institution and across different institutions. In my view, the national systematization of academic salaries is a good thing, and this is an area in which the USA might learn from the UK. It is mainly for the good, I think, that the UK has not emulated the USA in awarding higher salaries to stars and in playing off colleagues against each other as they search for higher and higher wages. This is not to say that academic salaries are by any means adequate in this country. They are surely not, even if one takes into account the different financial circumstances of educational institutions in the UK and the USA. The salary scales here desperately need to be adjusted upward.

Classics in the UK is an intellectually lively discipline, and it is producing some of the finest students and the most challenging scholarship anywhere. I offer these remarks as practical suggestions, and not as complaints, for the improvement of a system that, by and large, works well.

PHIROZE VASUNIA
UNIVERSITY OF READING

CLASSICS AT UK UNIVERSITIES, 2005–6
STATISTICS

As in previous years, data are divided into (a) 'traditional' Classics courses, such as BA Classics, Greek, or Latin, (b) 'modern' variants such as Classical Civilization, Classical Studies, Ancient History, and Classical Art and Archaeology, and (c) 'others' such as Combined Honours, supplementary students, and non-honours students. The grouping of data is as in previous years, with figures from the Open University fully integrated except where noted.

Trends in student numbers continue, with modest increases in all categories, save degrees in Joint Honours, which show a low, secular decline. It may be that these figures do not reveal the true picture. At my own university, borrowing of papers from other courses makes it possible to construct what are effectively joint degrees between Classics and Modern Languages, Oriental Studies and History; though these combinations are never referred to as 'joint degrees'. The point was made at a recent meeting of the Standing Committee that the time might have come to rethink our categories of courses to reflect more accurately what actually happens in Classics departments. I note, incidentally, that the graph that is Figure 1 is traditionally headed 'Traditional versus Modern Classics', though surely no competition is intended. Accordingly, next year's Statistical Officer might consult departments concerning ways in which returns might helpfully be altered. This could include, for example, some indication of involvement in Continuing Education in Classical subjects, which remains a grey area. Two considerations should be borne in mind. It is important not to add to the complexity of compiling departmental returns. Even in my short time as Statistical Officer it has proved increasingly difficult to achieve a 100% return. No criticism is intended; this is simply a reflection of the increasing burden of administration from which we all suffer. It is also desirable that we are able to preserve broad continuity in our statistics over time, making comparisons between years meaningful.

It seems likely that the statistics compiled by CUCD have found their way into the wider public domain. Recent correspondence in the national press over problems over the teaching of Modern Languages at school and university has drawn comparison with the success of Classics in maintaining recruitment.

PAUL MILLETT
DOWNING COLLEGE, CAMBRIDGE

Table A: Overview

	Honours students (SH + JH)				All students (incl. Other)								Staff FTE	Student-staff ratio	1st yr honours (SH + JH)			
					excluding OU				including OU				excluding OU	excluding OU				
	no.	FTE	% change FTE	index (1995–6 = 100)	no.	FTE	% change FTE	index (1995–6 = 100)	no.	FTE	% change FTE	index (1997–8 = 100)			no.	FTE	% change FTE	index (1995–6 = 100)
1995–6	5,606	3,804		100	9,356	5,317		100					361	14.7	2,152	1,288	-3.9	100
1996–7	5,647	3,812	0.2	100	9,269	5,095	-4.2	96					365	14.0	2,122	1,272	-1.3	99
1997–8	5,762	4,006	5.1	105	9,219	5,289	3.8	99	16,616	6,252		100	356	14.9	2,109	1,351	6.2	105
1998–9	5,610	3,898	-2.7	102	9,878	5,148	-2.7	97	16,610	6,119	-2.1	98	351	14.7	2,071	1,291	-4.4	100
1999–00	5,869	4,121	5.7	108	8,882	5,233	1.7	98	18,922	6,961	13.8	111	343	15.3	2,275	1,405	8.9	109
2000–1	5,499	3,803	-7.7	100	8,665	4,996	-4.5	94	16,634	6,475	-7.0	104	360	13.9	2,125	1,362	-3.1	106
2001–2	5,673	3,858	1.4	101	8,244	4,549	-8.9	86	18,786	6,195	-4.3	99	361	12.6	2,293	1,399	2.7	109
2002–3	5,571	4,225	9.5	111	8,577	5,016	10.3	94	17,507	6,394	3.2	102	377	13.3	2,177	1,585	13.3	123
2003–4	5,854	4,527	7.1	119	8,399	5,129	2.2	96	17,866	6,460	1.0	103	388	13.2	2,302	1,552	-2.1	121
2004–5	5,834	4,571	1.0	120	8,475	5,325	3.8	100	16,986	6,349	-1.7	102	371	14.3	2,205	1,599	3.0	124
2005–6	6,186	4,868	6.5	141	8,937	5,500	3.3	103	17,448	6,524	2.8	104	396	13.9	2,234	1,684	5.3	131

42

Table B: Single and Joint Honours

	'Traditional classics' (Classics, Greek, Latin)								'Modern classics' (Class. Studs, Anc. Hist., Art/Arch.)							
	single honours				joint honours				single honours				joint honours			
	no.	FTE	% change FTE	index (1995–6 = 100)	no.	FTE	% change FTE	index (1995–6 = 100)	no.	FTE	% change FTE	index (1995–6 = 100)	no.	FTE	% change FTE	index (1995–6 = 100)
		trad SH				trad JH				mod SH				mod JH		
1995–6	1,234	1,162	-3.0	100	323	139	-6.1	100	2,099	1,661	-12.0	100	1,950	842	8.4	100
1996–7	1,165	1,098	-5.5	94	299	130	-6.7	93	2,011	1,704	2.6	103	2,172	881	4.6	105
1997–8	1,243	1,159	5.5	100	263	118	-9.2	85	2,207	1,822	6.9	110	2,049	908	3.1	108
1998–9	1,241	1,181	2.0	102	333	155	31.6	112	2,001	1,711	-6.1	103	2,035	851	-6.2	101
1999–00	1,178	1,074	-9.1	92	298	120	-22.9	86	2,375	2,036	19.0	123	2,018	892	4.7	106
2000–1	1,109	1,019	-5.1	88	219	97	-19.0	70	2,068	1,824	-10.4	110	2,103	863	-3.2	102
2001–2	1,082	968	-5.0	83	265	127	31.3	91	2,363	1,958	7.4	118	1,963	805	-6.7	96
2002–3	1,108	1,022	5.6	88	238	120	-5.3	87	2,525	2,255	15.1	136	1,700	828	2.9	98
2003–4	1,362	1,150	12.5	99	221	126	5.1	91	2,582	2,363	4.8	142	1,689	879	6.1	104
2004–5	1,482	1,225	6.5	105	232	114	-9.4	82	2,518	2,424	2.6	146	1,602	809	-8.0	96
2005–6	1,624	1,300	6.2	112	200	96	-16.2	69	2,792	2,576	6.3	155	1,571	896	10.9	106

Table C: All Students

| | 'Traditional Classics' | | | | | | | | 'Modern Classics' | | | |
| | Classics | | Greek | | Latin | | Class. | | Anc. Hist. | | Class. Art/Arch. | |
	no.	FTE	no.	FTE	no.	FTE	no.	FTE	no.	FTE	no.	FTE
SINGLE HONOURS												
2000–1	1,039	953	17	17	53	49	1,179	1,067	791	685	98	72
2001–2	1,011	920	19	13	52	35	1,235	1,094	1,006	781	122	83
2002–3	1,025	951	29	24	54	47	1,281	1,172	1,072	935	172	148
2003–4	1,237	1,059	41	31	84	59	1,179	1,118	1,296	1,090	106	146
2004–5	1,346	1,133	44	31	92	60	1,280	1,214	1,109	1,056	129	154
2005–6	1,462	1,200	68	46	94	55	1,444	1,313	1,224	1,110	124	153
JOINT HONOURS												
2000–1	53	26	30	12	136	59	627	292	1,180	498	216	72
2001–2	11	7	43	19	211	101	694	306	913	403	357	95
2002–3	34	20	19	8	185	92	461	245	1,149	532	90	51
2003–4	64	36	12	6	145	84	522	313	1,046	512	122	72
2004–5	72	36	17	8	143	70	477	260	1,063	498	62	51
2005–6	66	32	33	14	101	49	539	310	975	552	57	35
OTHER												
2000–1	9	7	48	20	105	33	7,449	1,355	318	75	140	38
2001–2	3	2	538	139	946	244	10,163	1,570	1,168	287	244	73
2002–3	3	2	809	208	742	206	8,778	1,366	1,236	303	368	85
2003–4	74	13	575	147	642	165	9,020	1,220	1,365	377	425	97
2004–5	7	4	633	151	637	162	8,100	924	1,452	441	323	97
2005–6	17	7	667	155	733	174	8,124	895	1,315	307	406	118
ALL												
2000–1	1,101	986	95	49	294	141	9,255	2,714	2,289	1,258	454	182
2001–2	1,025	928	600	171	1,209	381	12,092	2,971	3,086	1,471	723	251
2002–3	1,062	973	857	240	981	345	10,519	2,784	3,457	1,769	630	284
2003–4	1,375	1,108	628	185	871	309	10,721	2,651	3,707	1,979	653	315
2004–5	1,425	1,173	694	191	872	292	9,857	2,397	3,624	1,994	514	302
2005–6	1,545	1,239	768	215	928	278	10,107	2,518	3,514	1,968	587	306

Figures in italics include Open University data.

Table D: Staff

| | Full-time | | | | Part-time | | | | Other | |
| | permanent | | temporary | | permanent | | temporary | | | |
	no.	FTE	no.	FTE	no.	FTE	no.	FTE	no.	FTE
2000–1	335	**316**	39	**37**	5	**3**	84	**25**	122	**34**
2001–2	332	**330**	47	**44**	12	**4**	74	**28**	156	**48**
2002–3	333	**323**	49	**49**	9	**20**	82	**37**	142	**29**
2003–4	327	**324**	41	**41**	12	**5**	75	**35**	148	**35**
2004–5	345	**342**	38	**39**	40	**19**	53	**15**	150	**56**

Summary 2005–6 (all staff)

	no.	FTE	% change	on leave (FTE)	effective (FTE)	% change
2000–1	585	**413**		53	**361**	
2001–2	621	**454**	9.8	68	**386**	7.0
2002–3	615	**458**	1.1	63	**395**	2.4
2003–4	602	**439**	-4.2	60	**379**	-4.1
2004–5	625	**471**	7.4	68	**404**	6.5

FTE since 1992–3

	no.	% change		no.	% change
1992–3	347		1999–00	343	-2.3
1993–4	352	1.3	2000–1	360	5.1
1994–5	379	7.6	2001–2	361	0.1
1995–6	361	-4.6	2002–3	386	7.0
1996–7	365	0.9	2003–4	395	2.4
1997–8	356	-2.3	2004–5	379	-4.1
1998–9	351	-1.5	2005–6	404	6.5

Figures exclude the Open University

Table E: Beginners' Languages

Undergraduates

	Greek			Latin		
	no.	FTE	% change	no.	FTE	% change
2000–1	626	134		1,206	270	
2001–2	1,052	278 *		1,398	359 *	
2002–3	983	259	-7	1,234	309	-14
2003–4	901	232	-11	1,228	321	4
2004–5	976	302	30	1,319	348	8
2005–6	1,015	269	-11	1,294	329	-5

Postgraduates

	Greek			Latin		
	no.	FTE	% change	no.	FTE	% change
2001–2	44	13		72	20	
2002–3	33	9	-29	41	15	-23
2003–4	33	6	-34	72	16	5
2004–5	55	14	128	81	16	-1
2005–6	53	16	18	78	20	26

Table F: Postgraduates

	Full-time	Part-time	Other (FTE = 0)	Total no.	**FTE**	*% change*
TAUGHT						
2000–1				319	**245**	
2001–2	240	183	11	434	**331**	
2002–3	246	242	8	496	**357**	*7.9*
2003–4	268	256	8.7	532.7	**373**	*4.4*
2004–5	277	222	9	508	**354**	*-5.1*
2005–6	315	208	10	533	**423**	*19.4*
RESEARCH						
2000–1				420	**363**	
2001–2	339	126	41	506	**393**	
2002–3	361	123	39	523	**410**	*4.2*
2003–4	388	157	14	559	**442**	*7.9*
2004–5	411	130.5	18	559.5	**482**	*9.0*
2005–6	432	107	14	553	**468**	*-2.8*

Fig. 1. FTE student numbers in UK for 'traditional' v. 'modern' classics, 1995–2006. (Source: CUCD.)

black = 'traditional', **grey = 'modern'**; **solid line = single honours (SH), broken = joint honours (JH)**