

In the press: Science and Public Policy

Evaluating universities using simple scientometric research output metrics:
Total citation counts per university for a retrospective seven year rolling sample

Bruce G Charlton & Peter Andras

ABSTRACT

We advocate a scientometric, top-down, and institution-based research assessment methodology that is based on total citations accumulated from all publications associated-with a specific university during the survey period. The exercise could be done every year using a rolling 7 year retrospective sample and should be performed by at least two independent auditors. Identification of elite 'revolutionary science' institutions could be accomplished using a metric derived from the distribution of science Nobel Prizes.

INTRODUCTION

The purpose of the UK Research Assessment Exercise (RAE) is to measure the quality of research in UK universities, with the aim of providing central government funding to support the long-term research capability of an institution.¹ At present, the UK RAE is a 'bottom-up' and discipline-based expert review process determined using a common set of information provided by each disciplinary unit within each university.

By contrast, we advocate a 'top-down' institution-based research assessment methodology based on total citations accumulated from all publications associated-with a specific university during the survey period. Such a survey could be done every year using a rolling 7 year retrospective sample and performed by at least two independent auditors. Identification of elite 'revolutionary science' institutions could be based-on a metric derived from the distribution of science Nobel Prizes.

THE SCIENTOMETRIC TOP-DOWN APPROACH TO METRICS-BASED RESEARCH EVALUATION

Our suggested usage of a RAE based on a metric of total citations from all publications associated-with a specific university is an example of top-down research evaluation, using a single 'macro'-level variable. By contrast, the current UK RAE is implicitly a bottom-up approach to research evaluation using an accumulation and average of many 'micro'-level evaluations.

This top-down approach to research evaluation derives from the discipline of scientometrics in which the evaluation of research is seen as a science in its own right.² Scientometric analysis is therefore typically performed by observers 'outside' the system being evaluated; in contrast to the bottom-up evaluations which are typically performed by 'peer review' of individuals with expertise in the discipline being evaluated. Since its origins in the 1940s, scientometrics has developed its own 'system language' including information selections and a distinctive lexicon and grammar.³ For this reason, the top-down scientometric procedures of research

evaluation are usually completely different—from the academic work that goes on within those individual academic disciplines that are being evaluated.

The current UK RAE is therefore (implicitly) an example of bottom-up research evaluation since it uses the internal evaluation procedures of many individual academic disciplines (approximately 70 disciplines for RAE 2008). These evaluations are performed by peer review panels of experts with the relevant disciplinary expertise, and the criteria of evaluation are very different between disciplines such as mathematics, biochemistry, English literature, social policy and the performing arts. To generate a metric for each university, these dissimilarly-derived evaluations are averaged. The validity of the summary RAE metric is therefore a bottom-up consequence of the validity of all the discipline-specific metrics which have gone into it, and of the averaging procedures.

By contrast, the validity of a top-down scientometric evaluation procedure is not a product of its constituent parts, but emerges as a consequence of its usage. The validity does not, therefore, depend upon discipline-level information; but on how well the metric performs as a measure. Scientometrics, as a science, therefore evolves by the usual scientific processes, such as making hypotheses and predictions which are tested by further observations.⁴

ANALOGY BETWEEN SCIENTOMETRICS AND MACROECONOMICS

The difference between top-down scientometric and bottom-up approaches to research evaluation is somewhat analogous to the difference between top-down macroeconomics and bottom-up microeconomics.⁵

Top-down macroeconomics examines total national economic activity in terms of variables such as taxes and interest rates; and national behaviours such as economic growth and inflation. Bottom-up microeconomics examines the behaviours of individual people and organizations in terms of variables such as money and time; and incentives such as pay and profit.

Microeconomic analysis therefore aspires to discover, understand and predict the actual economic incentives directly experienced by individuals and organizations (even when individuals or organizations may not be consciously aware of these incentives). By contrast, macroeconomics uses variables which have 'proved themselves' as useful at the national level (or, at least, more useful than the known alternatives) in terms of tasks such as monitoring, predicting and controlling the national economy.

The validity or usefulness of a macroeconomic variable is not necessarily challenged by critiques from a microeconomic perspective. For example, although raising the central bank interest rate by a quarter of a percent is a recognized (and usually effective) method of reducing inflation, this is hard to understand from the perspective of the incentives of individual citizens who would probably not be able to detect the consequences of such a small alteration in their personal finances – given the number of other and larger influences on personal finance. And the Gross Domestic Product (GDP) metric generally performs well as a longitudinal and internationally-comparative measure of the size of a national economy, yet GDP can easily be criticized from a microeconomic perspective as both incomplete and biased.

Similarly, our proposed research evaluation metric of the total annual citations associated with a university can legitimately be criticized from a micro perspective as incomplete and biased.⁶ For example citations in books may be omitted, and the total

citations metric is often dominated within a university by very high citation rates in relatively few fields within the natural sciences. Nonetheless, while correct at the micro level, none of these or similar criticisms are necessarily decisive at the macro level.

What matters at the macro level is whether the total citations metric performs well in the function for which it is intended.

THE PRIMARY FUNCTION OF RESEARCH EVALUATION IS TO MEASURE SCIENCE IN A WAY WHICH ALLOWS LONGITUDINAL AND INTERNATIONAL COMPARISONS

Although the UK RAE currently presents itself as a method of evaluating the whole of academic research activity, we believe that the implicit motivation of research evaluation is focused upon ‘scientific’ research – that is the mathematical and natural sciences, and the quantitative social sciences such as economics. And the most useful type of RAE would enable scientific research to be compared longitudinally and internationally.

The current UK RAE uses a highly complex, non-verifiable, un-checkable, evolving, bottom-up, discipline-based, peer review process which lacks transparency. So the RAE results cannot legitimately be used to track longitudinal changes of individual universities, nor can it be used to measure the relative strength of UK universities internationally (either individually or as a national whole). Furthermore, the RAE results are not checkable or replicable, even in principle. We regard these as very significant criticisms of any discipline-based peer review system of research evaluation when used for measuring institutional scientific research performance.

We would argue that it should explicitly be accepted that the primary implicit focus of research evaluation is scientific. Indeed, the field of scientometrics originated in the realm of evaluating scientific research during the period of rapid post-1945 expansion of science (including the industrial-organization of research) and of the large-scale government funding of science.⁷ Science research metrics have a generally-good track record, and the scientometric research of organizations such as the Institute for Scientific Information (ISI – now Thomson Scientific) is heavily used at many levels within science - from international, through national to local - where it exerts a significant influence upon the conduct of science policy including funding.

In a nutshell, science research metrics matter because scientific research matters to many powerful individuals and groups outside of science. By contrast, there are few compelling reasons for wishing to measure non-scientific research performance using metrics. To be blunt, non-scientific research is believed (by those outside it) to lack to the critical national importance of science. Scientific research involves vast funding and large groupings of organized personnel; but this does not apply to non-scientific research. Scientific research is considered to influence national prosperity, security and standard of living; but the same is not widely believed to be the case for non-scientific areas of research.

Furthermore, research performance metrics in the Arts and Humanities are generally felt by insiders to be much less valid than in scientific fields. In sum, it can be argued that non-scientific disciplines neither need top-down research metrics, nor do academic specialists in non-scientific disciplines accord such metrics the respect which they are given in the natural sciences.

The primary focus of research metrics is implicitly to measure and evaluate science, therefore we suggest that research assessment should be optimized for this function. A research metric should have qualities such as simplicity, transparency, objectivity, replicability, precision and sensitivity. In particular, research evaluation should be done using metrics which allow for both longitudinal (over time) and international (across space) comparisons. When a metric has these qualities it enables scientometrics to develop scientifically, because rival metrics can be checked to see if they perform adequately, and compared to see which metric performs the best overall.

THE SUPERIORITY OF RESEARCH-OUTPUT METRICS

Research metrics can be based on input and output variables. Input variables are the resources which generate research, and these include research workers with varying levels of training and skill, institutional infrastructure, and research funding. Output metrics include publications, estimates of the quality of these publications such as the Impact Factor of journals in which publications appear, and citations generated by these publications (which are a measure of the actual impact of publications).

One suggestion from the Higher Education Funding Council for England (HEFCE) has been that future RAE allocation should be based on input metrics such as external research income (including research council funding and other sources of funding from charities and foundations).⁸ For instance the proportion of RAE money a university received might be the same as the proportion of external funding it had won in competition with other universities. Other possible input metrics suggested have included the number of research staff or PhD students.

However, we argue that research metrics for an RAE should primarily reflect research output, and not inputs. While it is true that research inputs and outputs for the UK 1997-2003 show a close statistical correlation at any given point in time (eg. Andras & Charlton, unpublished observations), the use of inputs such as research council income as an RAE funding metric provides an incentive for consuming resources rather than generating research, and over-time this would probably lead to greater inefficiency.

A further serious problem with basing an RAE on research grant income is that funding agencies make their decisions on the basis of expert peer review – which is (as described above) a relatively non-transparent and unaccountable method of evaluation. Therefore, the use of input measures would tend to prevent the RAE results being used in longitudinal and international studies of scientific research performance: put briefly, a metric based on income is not well-suited to scientific validation by scientometric methods.

CORRELATION BETWEEN CITATIONS AND THE SJTU RANKING

We favour an RAE based on research output metrics, specifically citations. The validity of citations as a measure of a university's research performance can be tested empirically by comparing it with independent measures of university research performance.

While rankings based on publication numbers and citations are closely correlated (see below), we would advocate the use of citations in preference to publications. An RAE based on simple publication counts would create a perverse incentive, being prone to manipulation by 'salami-slicing' research output into minimum-publishable-units

(MPUs) in order to achieve maximum numbers. By contrast, citations are harder to manipulate. Salami-slicing would be discouraged by a citation-based metric because MPUs are of the lowest-possible quality while still being publishable – and such publications would tend to attract very few citations.

A further possible manipulation of citations might be individual or small group self-citations. These could potentially be screened-out using computerized methods. However, since self-citation is a potentially important aspect of the scientific process,⁹ the best way to decide whether to make adjustments or corrections to the method of citation counting would be comparative – the performance of a total metric (including self-cites) should be compared (longitudinally and internationally) with a metric screening-out self-citations to see which metric yields the best results overall.

This experimental attitude is indeed the general approach we advocate for resolving most methodological debates concerning top-down metrics: the best answer to questions of validity is to run a comparison of metrics over time, and see which one seems most useful overall. Where differences between metrics are small, the simpler metric should be favoured as being more transparent and easy to replicate.

METHODS

We recently analyzed thirty years of publications and citations from 47 UK universities which were universities before 1975 and which did not change their names between 1975 – 2004. The database was the ISI Web of Knowledge (WoK) which pools three specialized indices: Science Citation Index (Sci Cit), Social Sciences Citation Index (Soc Sci), and Arts and Humanities Citation Index (A&H).

Excluded were Cardiff University (due to name changes) and University of Manchester Institute of Science and Technology (UMIST) (which has now joined University of Manchester). Many existing London colleges (Imperial, UCL, Kings, Royal Holloway) have during this time merged with smaller colleges, medical schools and other research institutes. We did not include separate medical schools (e.g. St George's Medical School) or research institutes (e.g. Cancer Research Institute).

All universities had unique search strings. For each considered university we counted the number of publications (all types of publications) for each year between 1975 – 2004 inclusive, for the combined three domains of WoK. We also counted for each university, and each year, the number of citations that the corresponding publications received between the time of their publication and the time of data collection (February 2006). When a publication was written by multiple authors from more than one included university, the publication was counted for each university to which at least one of its authors belonged.

The rankings of both publications and citations were very similar (Table 1). Furthermore, rankings for publications and citations were also predictive of ranking in the authoritative Shanghai Jiao Tong University (SJTU) Academic Ranking of World Universities published in 2005¹⁰ with citations providing a slightly-closer correlation than publication numbers (Figures 1 & 2).

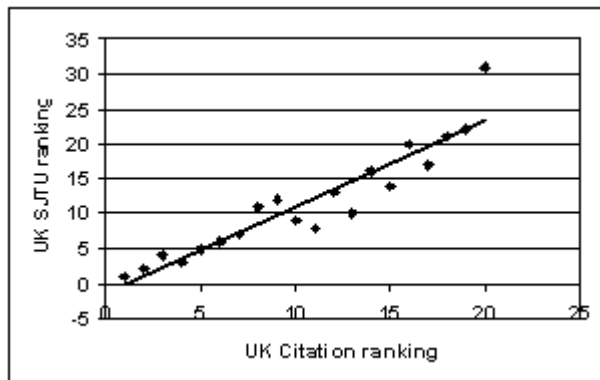
The SJTU ranking sets out to measure excellence in (implicitly) scientific research; and it depends on a weighted scale including Nobel Prize/ Fields Medal (10% for Alumni and 20% for staff), Number of ISI highly-cited researchers (20%), number of articles published in Nature and Science (20%), number of articles in Web of Science (20%) and a 10 % adjustment to control (partly) for the size of institution. Since the SJTU score derives from a different set of data than citations (with the partial

exception that citation counts are used to choose ISI highly-cited researchers) this provides an independent validation of citation counting as a measure of research excellence.

Table 1: Rank based on number of WoK number of citations/ number of publications for 5 years 2000-2004 - (SJTU UK Rank)

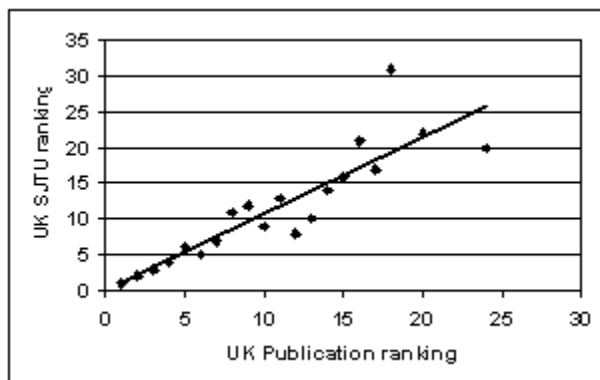
1/ 1 (1)	Cambridge
2/ 2 (2)	Oxford
3/ 4 (4)	UCL
4/ 3 (3)	Imperial College, London
5/ 6 (5)	Edinburgh
6/ 5 (6)	Manchester
7/ 7 (7)	Bristol
8/ 8 (11)	Birmingham
9/ 9 (12-15)	Glasgow
10/ 10 (9)	Kings College, London
11/ 12 (8)	Sheffield
12/ 11 (12-15)	Leeds
13/ 13 (10)	Nottingham
14/ 15 (16-19)	Southampton
15/ 14 (12-15)	Liverpool
16/ 24 (20-20)	Dundee
17/17 (16-19)	Leicester
18/ 16 (20-30)	Newcastle
19/ 20 (20-30)	Durham
20/ 18 (>30)	Aberdeen

Figure 1: SJTU 2005 Ranking for UK universities plotted against ranking for total WoK citations 2000-2004



Spearman Rank Correlation: $R=0.967$, $p<0.001$.

Figure 2: SJTU 2005 Ranking for UK universities plotted against ranking for total WoK publications 2000-2004



Spearman Rank Correlation: $R=0.949$, $p<0.001$.

Correlations between citations and the SJTU ranking are especially close at the top of the rankings where there are also large differences between the universities in terms of metrics – correlations are less impressive when the distribution curve is flatter and differences between universities are smaller. Another factor is that the SJTU ranking exhibits increasing levels of statistical ‘noise’ in lower ranks probably due to small-number effects. Only a few elite universities have numerous Nobel/ Fields/ highly-cited researchers; and when there are few individuals in these categories just one more or less can make a significant difference to the SJTU rank.

Consistent with this UK data is that Harvard University (SJTU number 1 in the world) is also the most productive and most cited university we have measured – generating both more publication and more citations per year than Cambridge and Oxford combined.¹¹ In general, this analysis seems to indicate that citation counts are correlated with an independent measure of the scientific research quality of UK universities.

PROPOSAL FOR A CITATION-BASED RAE – THE COMPETITIVE PRINCIPLE FOR EVALUATION

In the light of the above analysis, we advocate replacing discipline-based expert review with an institution-based research output metric: total citations accumulated from all publications associated with a specific university during the survey period.

Citations could be surveyed by at least two independent auditors to be chosen by competitive tendering, and using databases such as the ISI Web of Knowledge¹² and Elsevier Scopus¹³. Indeed, Thomson Scientific (ISI) and Elsevier might be suitable organizations for commissioning to perform citation analyses for the RAE, since they have access to raw data, are in competition, and therefore each seem likely to be motivated to do a good job. This competitive system of research assessment fits with our advocacy of a scientometric approach. The best methods for research metrical analysis will tend to emerge with time and over repeated cycles of evaluation in terms of attributes such as superior applicability, precision, predictive power, simplicity, transparency, cheapness etc.

Simple output metrics are transparent, clear and cheap; and can be measured using independent external expertise, without involvement of those being measured. The objectivity and transparency of a simple citation metric would enable universities to manage strategically, since they can calculate for themselves in advance the approximate level of future RAE funding derived from measuring their own citation growth compared with other relevant institutions.

Due to its cheapness and simplicity this kind of top-down RAE could be performed every year. But clearly it is an advantage to use several years of accumulated citations in order to obtain a valid and precise measurement. We therefore suggest using a rolling 7 year retrospective sample of accumulated citations – 7 years being chosen on the basis that that it is a reasonable approximation to the timescale of scientific activity (5 years is probably too short while 10 years is probably too long)¹⁴. The use of a rolling retrospective sample will also smooth-out year by year changes, and avoid sudden large reductions in annual funding which could prove needlessly disruptive to institutions.

The raw citation counts would be used to create a rank ordering of universities which may be fitted to a funding formula curve. The shape of the funding curve would be determined by strategic research priorities.

What would be the likely incentives created by this total citation metric? Probably, since total citation counts are heavily weighted in favour of natural sciences, the main incentive would be for universities to compete in attracting the most-cited scientific research teams in the leading branches of the natural sciences and quantitative social sciences. This would raise the cost of the most-cited research teams, probably improving the pay, support and conditions of group members, and further increasing competition to succeed in highly-cited fields. We believe such changes would, on the whole, be beneficial to scientific research.

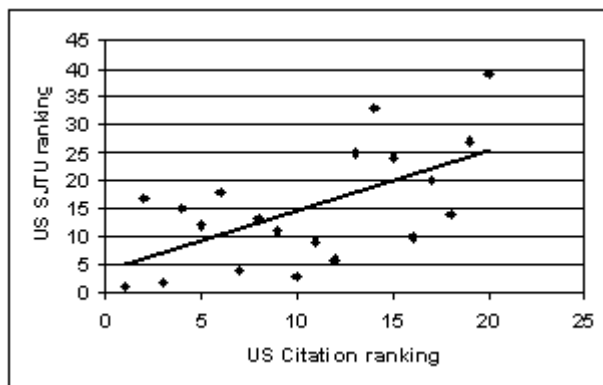
ELITE 'REVOLUTIONARY SCIENCE' UNIVERSITIES

While in the UK the SJTU and citation rankings are very similar, the top US universities display some significant dissociations between the SJTU rankings and rankings by total WoK citations (ie including Sci Cit, Soc Sci and A&H) (Table 2).

Table 2 - Top 20 US universities ranked according to SJTU 2005, and according to number of WoK citations 2000-2004.

SJTU - 2005	Citations – 2000-2004
1. Harvard	1. Harvard
2. Stanford	2. Johns Hopkins
3. U California Berkeley	3. Stanford
4. MIT	4. U of Washington, Seattle
5. CalTech	5. UCLA
6. Columbia	6. U Michigan, Ann Arbor
7. Princeton	7. MIT
8. Chicago	8. U Pennsylvania
9. Yale	9. U California, San Diego
10. Cornell	10. U California, Berkeley
11. U California, San Diego	11. Yale
12. UCLA	12. Columbia
13. U of Pennsylvania	13. U of Minnesota
14. U of Wisconsin, Madison	14. U of Pittsburgh
15. U of Washington, Seattle	15. Duke
16. U California, San Francisco	16. Cornell
17. Johns Hopkins	17. Washington U , St Louis
18. U of Michigan, Ann Arbor	18. U of Wisconsin, Madison
19. U of Illinois, Urbana Champaign	19. U of Colorado
20. Washington U, St Louis	20. U of North Carolina, Chapel Hill

Figure 3: SJTU 2005 ranking of US universities plotted against ranking for total WoK citations 2000-2004.



Spearman Rank Correlation: $R=0.564$, $p=0.01$.

Some universities are significantly higher in the SJTU and are placed significantly below the line on Figure 3: For example, Berkeley is 3rd in SJTU and 10th in citations; Princeton is 7th and Chicago 8th in SJTU and they are not even in the top 20 for citations. Conversely, other universities are above the line on Figure 3, such as Johns Hopkins which is 2nd for citations and only 17th in SJTU and the University of Washington, Seattle is 4th in citations and 15th in SJTU.

This pattern is probably influenced by subject mix, since among the top 10 SJTU universities Berkeley and Princeton do not have medical schools, and Chicago only a small one. Conversely, Johns Hopkins and University of Washington, Seattle are major medical research centres. There is also a trend for private universities to rank relatively higher in the SJTU (9 out of the top 10) than in the citations ranking (5 out of the top ten).

Since the US universities include the best and most complexly differentiated in the world, this pattern may represent a specialization into Kuhnian categories of 'revolutionary' and 'normal' science (especially biomedical science) among the most productive US universities.¹⁵ The revolutionary science top universities (such as Berkeley, Princeton and Chicago) are most likely to gain Nobel Prizes/ Fields medals, publish in Nature and Science, and contain disproportionately large numbers of very highly-cited researchers.

The more 'normal science' top universities (eg. Johns Hopkins, U Washington Seattle, UCLA and U Michigan Ann Arbor) may be focused more on large scale biomedical research. In its Kuhnian sense, 'normal' science entails activities such as checking and incrementally-extending existing knowledge using established techniques, and improving existing research methods. However, it should be noted that by comparison with other countries than the US, even these more 'normal science' top-cited universities are high achievers in the indices of revolutionary science such as Nobel Prizes, ISI highly-cited academics and Nature and Science publications.

Our conclusion is that citation metrics probably do not always distinguish the highest levels of excellence in revolutionary science from less innovative but highly productive 'normal science'. Different metrics would be required for this purpose, and we suggest the use of Nobel Prizes.

INTERNATIONAL BENCHMARKING OF REVOLUTIONARY SCIENCE USING NOBEL PRIZES.

Some of the components of the SJTU metrics are undesirable as measures of revolutionary science. Nature and Science are weekly-published commercial journals which contain mostly 'scientifically-fashionable' papers expected to gather large numbers of citations; and they have an increasing role in generating media discussion of scientific issues, which probably interferes with their status as journals of record. The inclusion of numbers of publications, and of highly-cited scientists, again fail to pick-out revolutionary science from the most successful examples of normal science.

We suggest that science Nobel Prizes are the most promising basis for a metric of high quality revolutionary science, since they are generally-regarded as having very high validity for honouring the highest level of science in their fields. This would especially be the case if the Prizes could be reformed to generate more laureates per year and to provide an official allocation of credit to institutions.¹⁶

At present, there are four science Nobel Prizes – in Physics, Chemistry, Medicine/ Physiology and Economics; each is awarded to 1-3 laureates per year generating 4-12

laureates annually. Individuals are honoured, and there is currently no official allocation of credit between the scientific institutions which have nurtured and supported the Nobel-prize-winning work. The small numbers of laureates mean that only a few elite universities are able to accumulate sufficient Nobel prizes over a short-enough recent timescale to provide a measure of research quality which is both precise and useful for future policy. Nonetheless, some preliminary institutional analysis is possible, which we present here.

The Nobel Foundation lists the affiliations of laureates at the time the prize is awarded over the past century since the Nobel Prizes began in 1901¹⁷ (many of these citations are many decades old and of doubtful contemporary relevance). A few elite universities emerge with 10 or more laureates: Harvard 31; Cambridge 18; CalTech 17, MIT 17; Stanford 17; Columbia 16; Berkeley 15; Chicago 15; Princeton 10, Oxford 10 (also the Rockefeller Institute and University, New York - for graduate students only - has 14 laureates).

For comparison we can look at the evaluations generated by the authors of Wikipedia.¹⁸ They have created a tabulation which simply counts the number of laureates associated with each university since 1901. The calculation uses four categories of association: graduate; attendee or researcher; faculty member before or at the time of the award (this category includes the Nobel foundation listing above); and faculty member after the award – so multiple institutions may receive credit for a single laureate. The results are broadly consistent with the more restricted official Nobel figures based on time of award, with a top rank of institutional associations as follows: Cambridge 83; Columbia 81; Chicago 79; Harvard 76; MIT 63; Berkeley 61; Stanford 50; Oxford 47. The high degree of convergence between these two, admittedly approximate, metrics suggests that the use of Nobel Prizes to measure ‘revolutionary science’ has considerable promise.

As an example of using Nobel Prizes in a national RAE, we have selected only those UK universities that have had affiliated faculty who received two or more Nobel Prizes in the past 50 years since 1947-2006. This picks-out just six potentially elite ‘revolutionary science’ universities: Cambridge 6; Oxford 4; Imperial College, London 4; University College, London 3; King's College, London 2 and Sussex 2. The metric would be further improved if credit was primarily awarded to institutions which had supported the specific research which led to the Nobel award. So, Nobel Prizes could potentially identify the UK elite of ‘revolutionary science’ universities, whose status might justify a separate and extra funding stream to that deriving from the total citation count.

This metric could be made official and more valuable by reforms to the Nobel Prize system.¹⁹ We suggest more than doubling the number of Nobel science laureates from a maximum of twelve to a minimum of twenty-four per year. If there were more laureates being created, this would enable a wider range of outstanding scientific work to be recognized and a more precise estimate of the relative ability of universities and other research institutions to generate the highest quality of research. If the maximum number of three laureates per category were awarded as a matter of course, this would increase the annual number to twelve; and this number could be further increased by new categories of Prize (such as mathematics and computing science) and also by increasing the number of laureates in the ‘biological’ category of medicine/ physiology, which is now the largest and most prestigious branch of world science.

Furthermore, when awarding the Prize the Nobel committee could officially allocate differential credit for this work between research institutions, enabling a more valid

metric to be developed.²⁰ By such means the Nobel Prize might widen its role from honouring outstanding individual achievement to include a role in measuring and supporting outstanding scientific institutions.

CONCLUSION

We suggest that the proper purpose of a government research assessment is to generate a rank ordering of universities in terms of their scientific research output. We also believe that the evaluation process should be maximally transparent and accountable.

These constraints imply that the RAE process should be based on objective measures which can be checked by third parties. In order to minimize distortion or corruption, and avoid interference with the work of universities, it is preferable that the RAE does not require the cooperation of researchers. It is also desirable that the RAE is as cheap and quick as is compatible with validity. This suggests that an RAE should be based upon top-down quantitative metrics, and we suggest total citations is the best candidate.

We suggest that an institution-based RAE, using total-citations calculated on a rolling 7 year retrospective sample, is a metric which fulfils these major criteria. The validity of this scientometric method could continually be monitored on the basis of longitudinal and international comparisons of research performance, including other metrics.

The identification of elite 'revolutionary science' institutions could potentially be derived from a metric based-on the distribution of science Nobel Prizes. A Nobel metric could be used to ensure that universities with a record of supporting a significant volume of the highest quality of scientific research would get specific financial support.

Acknowledgement: thanks to Claire Donovan and Andrew Oswald for comments which helped improve this paper.

1. rae 2008 – Research Assessment Exercise. www.rae.ac.uk, accessed 16 Nov 2006.
2. Vannevar Bush, 'As we may think' [Atlantic Monthly](#), (July 1945). www.ps.uni-sb.de/~duchier/pub/vbush/vbush-all.shtml, accessed 15 November 2006; DJ de Solla Price, [Little science, big science - and beyond](#), (New York: Columbia University Press, 1986); Eugene Garfield, [Essays of an information scientist](#), (Philadelphia: ISI Press, 1977), 313-318; Loet Leydesdorff, 'The Evaluation of Research and the Evolution of Science Indicators', [Current Science](#), 89, (2005), 1510-1517.
3. Niklas Luhmann [Social Systems](#), (Cambridge, MA: Harvard University Press, 1995); Bruce Charlton and Peter Andras, [The modernization imperative](#), (Exeter, UK: Imprint Academic, 2003).
4. David L Hull, [Science as a process](#), (Chicago: Chicago University Press, 1988); John Ziman, [Real science](#), (Cambridge, UK: Cambridge University Press, 2000).
5. N Gregory Mankiw, [Principles of economics](#), (Mason, US: South Western College Publishers, 2003); Matthew Bishop, [Essential Economics](#), (London: Profile Books, 2004).
6. L. Butler, G. Laudel and C. Donovan, 'Quantitative indicators for research assessment', [Research Evaluation and Policy Project REPP Discussion Paper 05/1](#), (Canberra, Australia: Australian National University, 2005).
7. Bush, de Solla Price, Garfield and Leydesdorff, [op. cit.](#) note 2.

8. HEFCE - Higher Education Funding Council for England, 'Reform of higher education research assessment and funding', [Www.hefce.ac.uk/research/assessment/reform](http://www.hefce.ac.uk/research/assessment/reform) , accessed 16 November 2006.
9. Hull, *op.cit.*U note 4.
10. Institute of Higher Education, Shanghai Jiao Tong University, China, 'Academic ranking of world universities 2005: top 500 world universities', [Http://ed.sjtu.edu.cn/rank/2005/ARWU2005TOP500list.htm](http://ed.sjtu.edu.cn/rank/2005/ARWU2005TOP500list.htm) accessed 16 November 2006.
11. Bruce G Charlton and Peter Andras, 'Best in the arts, catching-up in science – what is the best future for Oxford?' *Oxford Magazine* 256, (2006), 25-6. [Www.hedweb.com/bgcharlton/oxmagarts](http://www.hedweb.com/bgcharlton/oxmagarts) , accessed 16 November 2006; Bruce G Charlton and Peter Andras, 'Oxbridge versus the 'Ivy League': 30 year citation trends', *Oxford Magazine* 255, (2006), 16-17. [Www.hedweb.com/bgcharlton/oxmagarts](http://www.hedweb.com/bgcharlton/oxmagarts) , accessed 16 November 2006.
12. Thomson Scientific, 'ISI Web of Knowledge' , [Http://portal.isiknowledge.com](http://portal.isiknowledge.com) accessed 16 November 2006.
13. Elsevier, 'Scopus' [Www.scopus.com/scopus](http://www.scopus.com/scopus) , accessed 16 November 2006.
14. Bruce G Charlton, 'Scientific life should be measured in seven year units', *Medical Hypotheses* , 66, (2006), 1051-2.
15. Thomas S Kuhn, *The structure of scientific revolutions*, (Chicago: Chicago University Press, 1970).
16. Bruce G Charlton, 'Why there should be more science Nobel prizes and laureates – and why proportionate credit should be awarded to institutions', *Medical Hypotheses* , in the press.
17. Nobelprize.org, 'Nobel laureates and universities', [Http://nobelprize.org/nobel_prizes/lists/universities.html](http://nobelprize.org/nobel_prizes/lists/universities.html) , accessed 16 November 2006
18. Wikipedia, 'Nobel prize laureates by university affiliation', [Http://en.wikipedia.org/wiki/Nobel_Prize_laureates_by_university_affiliation](http://en.wikipedia.org/wiki/Nobel_Prize_laureates_by_university_affiliation), accessed 7 Nov 2006.
19. Charlton, *op. cit.* note 16.
20. *Ibid.*