

Policy for ongoing evaluation of research output quality

Background

The current University Research Strategy includes two KPIs related to output quality: performance at next REF, and World citation impact metrics. Both of these are lagging measures – they reflect judgements of outputs made sometime after they have been published. Additional issues are that individual output grades from REF are never revealed, and citation metrics cover only some subjects and are highly subject-specific. Evaluating the quality of research outputs on an ongoing basis is fundamental to embedding a culture of improving output quality across the University, by ensuring that researchers at all career stages have a clear idea of what 2*, 3* and 4* outputs look like. Additional benefits are that:

- staff are further encouraged to produce high quality outputs rather than focus on volume; and
- outputs can be assessed at various stages in the writing process (e.g. pre-submission review), providing opportunities for increasing the quality (and potentially originality and significance) of the final work; and
- a robust and ongoing process of evaluating outputs quality would provide a transparent, fair and externally calibrated grading for use in the annual PDD, as well as promotion cases.

Outputs evaluation is also part of the essential preparations for REF2021.

Following discussions at Research Strategy, Policy and Performance Committee and College's Research Committees, the following processes have been approved.

Internal Processes

1. Initial training in output grading should be offered to staff via departments, utilising expertise from REF2014 panellists whenever possible. Staff should self-evaluate items to gain experience, but only grades agreed by internal panels and from external reviewers should be reported and retained by departments.
2. A department/UoA panel will evaluate research outputs by reading and grading them, following a process similar to that used for REF¹. Outputs can be graded at pre- or post-submission stage, but final grades should be based on published versions.
3. A 12 point scale is required, the same as that used for the REF Publications Fund. Appendix 1 gives an example of an output grading form. Each College should agree a standard form, as close to Appendix 1 as possible.
4. Departments must retain the evaluation reports (preferably in a specific area on the departmental X-drive) and, once the IRIS has the capability, ensure that internal and external grades are entered into IRIS.
5. Reviewing should take place on an ongoing basis, with each department/UoA reporting via an annual census, led by College Research Committees and reporting to RSPPC. It is suggested that annual reports are required in the autumn each year.
6. A standard reporting pro forma should be used (Appendix 2), requesting information on numbers of outputs reviewed, grades, percentage of staff taking part, panel members and working methods etc., and a brief commentary. The report should provide both an annual update, and a cumulative picture covering the whole REF period (i.e. 1 January 2014 onwards).
7. Open access publishing is of increasing importance both for REF and research funders. As well as reporting on output quality, the annual report should provide information on the proportion of published items available via open access (generally peer-reviewed conference proceedings and journal articles, but also some books and chapters), and hence eligible for REF and meeting research funder requirements.
8. Staff should submit their best outputs for evaluation, up to a maximum number per year. This will vary between disciplines, but a maximum value might be four.
9. For the 2017 reviews, a higher number would be allowed to ensure evaluation of all the high

This version February 2018

quality outputs published since 1 January 2014 (start of the REF census period).

10. One model for the panel is that three senior/experienced researchers, reflecting a spectrum of expertise across a subject area (similar in scope to a UoA), evaluate the papers. Other models, based on previous REF preparations, may also be appropriate.

¹A note on using REF-like criteria for evaluating the quality of research outputs: although measuring quality in the way suggested seems to be directly linked to REF, which is separate from the University research strategy, the criteria upon which REF grades are based are more broadly recognized as measures of research quality (e.g. Ellison 2002: Evolving standards for academic publishing: A q-r theory. *Journal of Political Economy* 110, 994-1034).

External Review

- a) A critical element of internal grading is that its robustness and accuracy is ensured through external calibration. External calibration is also required to counter the inevitable internal pressures to inflate grades.
- b) External grading should be sought on a mix of outputs: some selected by internal panels and some selected randomly (by RED).
- c) External reviews will be funded through the REF Publications fund, using reviewers nominated to RED. Around one third of potential REF outputs should be externally graded, with focus primarily on 4* and 3*/4* outputs.
- d) At least some of the external reviewers should have REF or RAE panel experience, although some reviewers might have other qualifications, such as extensive journal editing experience.
- e) Where there is significant disagreement about grades – either from internal review or across internal and external review – items may be sent to more than one external reviewer.

Recommendation 1: That the above policy be rolled out for the first annual review in 2017-18, which will cover outputs published from 1 January 2014 to 31 December 2017.

Research Strategy KPIs

As indicated above, the current KPIs for outputs do not help to improve or measure output quality at an appropriate time. A metric for an annual research output KPI might vary between colleges, or even between departments, to take into account differences in the nature of outputs produced and the research culture. It should be primarily designed to meet the need to have a meaningful KPI for research quality in University and College research strategies, but also serves other purposes (see above).

Recommendation 2: REF2021 will be based on an average of 2.5 outputs per FTE, (published from 1 January 2014 onwards), so an appropriate KPI for departments might be the cumulative number of accepted/published outputs graded at 4 and 3* on the annual census date as a percentage of total outputs required (T&R FTE x 2.5).*

Measuring research quality will add to the administrative burden, but would address concerns that the focus on income and award KPIs in University and College research strategies will change behaviour (which it is partly designed to do) in a way that de-emphasizes research quality. Research quality (especially outputs) feeds directly into external perceptions of reputation, and more directly into success in obtaining external research funding (track record is important).

Appendix 1: Example internal output grading form

Grading should be based on evidence of, or potential for, the characteristics below:

Authors:	
Output title and category (e.g. book, journal article etc.):	
Indicate all those that apply:	
Originality	research is leading or at the forefront of the research area (indicative of 4*)
	great novelty in developing new thinking, new techniques or novel results (indicative of 4*)
	develops new paradigms or fundamental new concepts for research (indicative of 4*)
	makes important contributions to the field at an international standard (indicative of 3*)
	involves incremental advances, which might include new knowledge which conforms with existing ideas and paradigms, or model calculations using established techniques or approaches (indicative of 2*)
Overall score for originality (4 maximum)	
Significance	agenda-setting (indicative of 4*)
	major influence on a research theme or field (indicative of 4*)
	contributes important knowledge/ideas/techniques, likely to have a lasting influence in field, but not necessarily leading to fundamental new concepts (indicative of 3*)
	provides useful knowledge and influences the field (indicative of 2*)
	useful but unlikely to have more than a minor influence in the field (indicative of 1*)
<i>changes in policy or practice (may not be applicable)</i>	Major (indicative of 4*)
	Significant (indicative of 3*)
	Some (indicative of 2*)
	Minor (indicative of 1*)
<i>influence on processes, production and management (may not be applicable)</i>	Major (indicative of 4*)
	Significant (indicative of 3*)
	Some (indicative of 2*)
	Minor (indicative of 1*)
<i>influence on user engagement (may not be applicable)</i>	Major (indicative of 4*)
	Significant (indicative of 3*)
	Some (indicative of 2*)
	Minor (indicative of 1*)
Overall score for significance (4 maximum)	
Rigour	<i>The purpose of the work is</i> articulated
	articulated reasonably well
	ambiguous/unclear
<i>Has the purpose been achieved?</i>	evidence presented
	some evidence
	unclear if achieved, or purpose not achieved
<i>Methodologies</i>	appropriate and rigorous
	lack rigour
Overall score for rigour (4 maximum)	
<i>Additional brief comments to justify grade</i>	Overall Grade total maximum = 12: 4 for originality 4 for significance 4 for rigour
Reviewer name:	Your expertise (delete as appropriate): high fairly high moderate fairly low low

This version February 2018

Appendix 2: Department/School Annual Output Quality Report Form

Name of Department/School:

Please provide the following information for the period XXX to YYY for the complete unit. Completed forms will be reviewed by College’ Research Committees who will report to Research Strategy, Policy & Performance Committee.

TYPE	No. of outputs published	INTERNAL REVIEW DATA						EXTERNAL REVIEW DATA					
		No. of outputs reviewed	% of outputs reviewed	Number at 4*	% at 4*	Number at 3*	% at 3*	Number reviewed	% of external reviews	Number at 4*	% at 4*	Number at 3*	% at 3*
Journal Article													
Book/ book chapter													
Peer-reviewed conference proceedings													
Other outputs													
Total													

1. How many REF eligible staff are in the School/department? How many of these had items (a) internally and (b) externally reviewed in the period?
2. Please give brief details of (a) how you grade items internally (including names of panel members); (b) how items are selected for external review; (c) how grades/comments are communicated to authors; (d) how reviews are stored and used internally; (e) any arrangements you have for mentoring and training around output quality. Please include here any other information about outputs and output quality that you think will be useful for the reviewing group.
3. How many (number and %) of the journal articles and conference proceedings are available as open access? Please explain how you monitor open access deposit and ensure items meet funding body and REF requirements.
4. What is the cumulative number (from 1 January 2014) of accepted/published outputs graded as 4* and the cumulative number of 3* as a percentage of total outputs required (T&R FTE x 2.5)?