

Hitting the QR sweet spot: will new REF2021 rules lead to a different kind of game-playing?



*Today marks 999 days until the expected deadline for submissions to REF 2021. Universities' preparations are already well under way, with additional guidance published last autumn in the form of new REF rules designed to reduce game-playing behaviours among institutions. However, as **Simon Kerridge** observes, the rule changes may have introduced, or rather enhanced, some hidden dangers around universities' FTE and impact submissions. Projections in funding allocation demonstrate why submitting institutions might be given pause for thought, with the driver for excluding staff to stay below an impact case study threshold possibly even higher than last time.*

The [Stern Review](#) sought to reduce game-playing behaviours among institutions, recommending the decoupling of outputs from individuals in order to reduce the potential divisiveness caused by non-selection for the Research Excellence Framework. These principles have informed certain REF2021 rules, which appear reasonable but after just a little analysis are revealed to have introduced, or rather enhanced, some hidden dangers.

The research impact component was introduced as part of REF2014 as a second sub-area with its own discrete elements – impact case studies and, at that time, an impact template too. Crucially, unlike with research outputs, additional chunks (case studies) were not needed for each additional person or FTE. This meant that whilst adding a person to a unit submission would mean providing more outputs, it would not necessarily require another impact case study – an extra one was only needed for every tenth FTE to be returned. As [shown elsewhere](#) (see also [slide 54 here](#)), this had a marked effect on the size of submissions. For submissions between 10 and 110 FTE, those of a size just under the requirement for an additional case study were around 7.5 times more common than those that just required an additional case study to be submitted (looking at one FTE either way). This was not the case for the 2008 Research Assessment Exercise when there was no case study element; the FTE distribution was more even (see Figure 1).

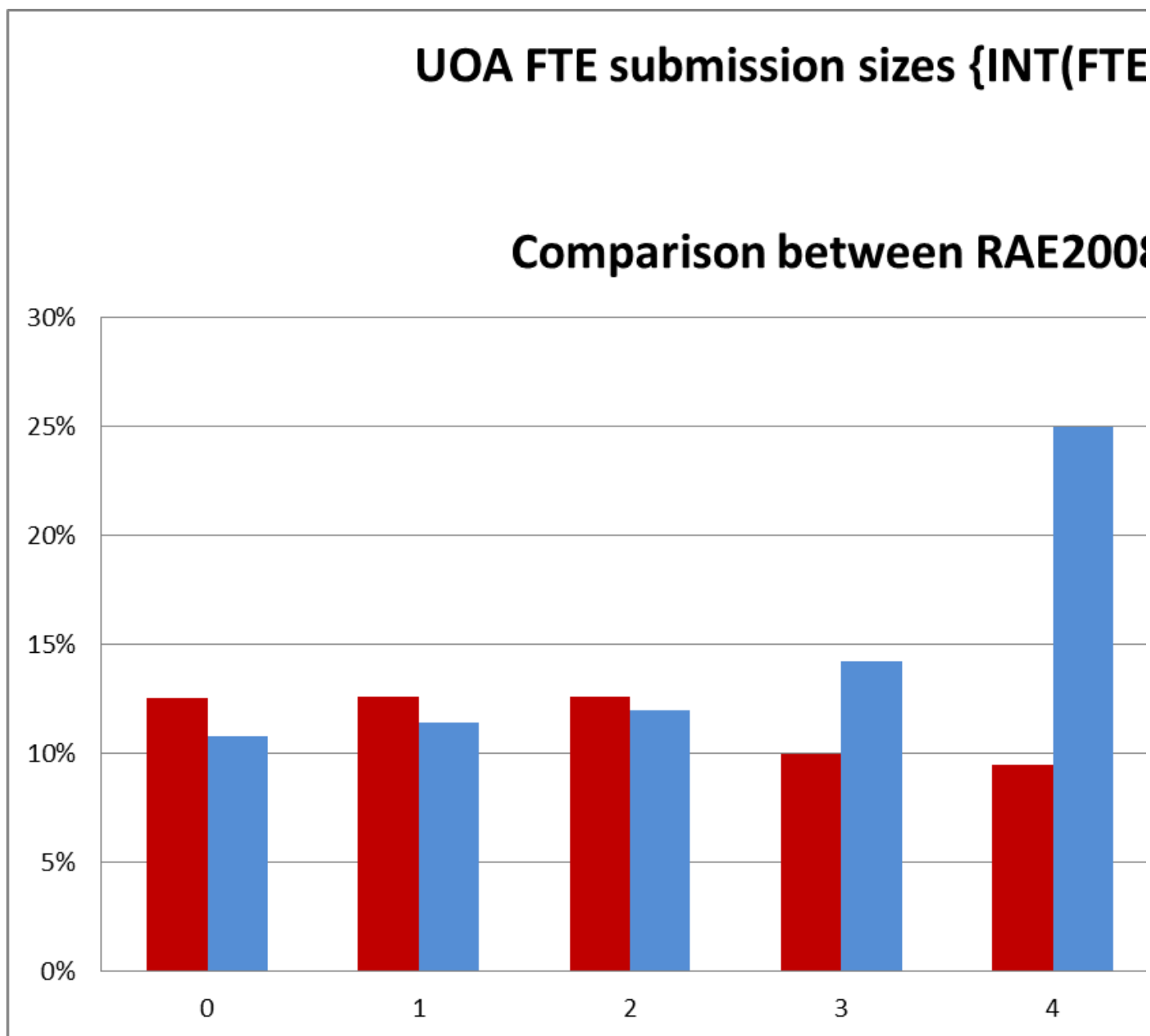


Figure 1: Distribution of submission sizes (modulo 10) for RAE2008 and REF2014. The largest single bar – at 4 on the x axis – is the point immediately below the threshold that would require submissions to provide an additional impact case study. Click to enlarge.

It seems that the potential impact (pun intended) of a weak impact case study lowering a quality profile and potentially decreasing the QR funding allocation was a powerful disincentive to increasing submission size beyond the next threshold. One can only imagine the effect on the 15th, 25th, 35th (and so on) members of staff who were not selected for submission.

It should be noted that, in terms of the overall QR funding allocation, an impact case study was on average worth 4.37 outputs in REF2014, a little more than one FTE's output contribution.

For REF2021 we have a [change of rules](#). Impact case studies are now worth 25% of the overall profile (up from 16% last time; a 56.25% rise), with outputs worth a little less at 60% (down from 65%; a 7.69% fall). However there is also a reduction in the numbers required of each – to compensate for the expected increase in FTE submissions as “[all staff with significant responsibility for research](#)” are to be returned. So instead of the (2014 average) of 3.67 outputs per FTE, the REF2021 average will be 2.5. This means the average value of an individual output will increase by 46.87%, which, when combined with the 7.69% fall mentioned above, gives a net overall increase in output value of 35.57%.

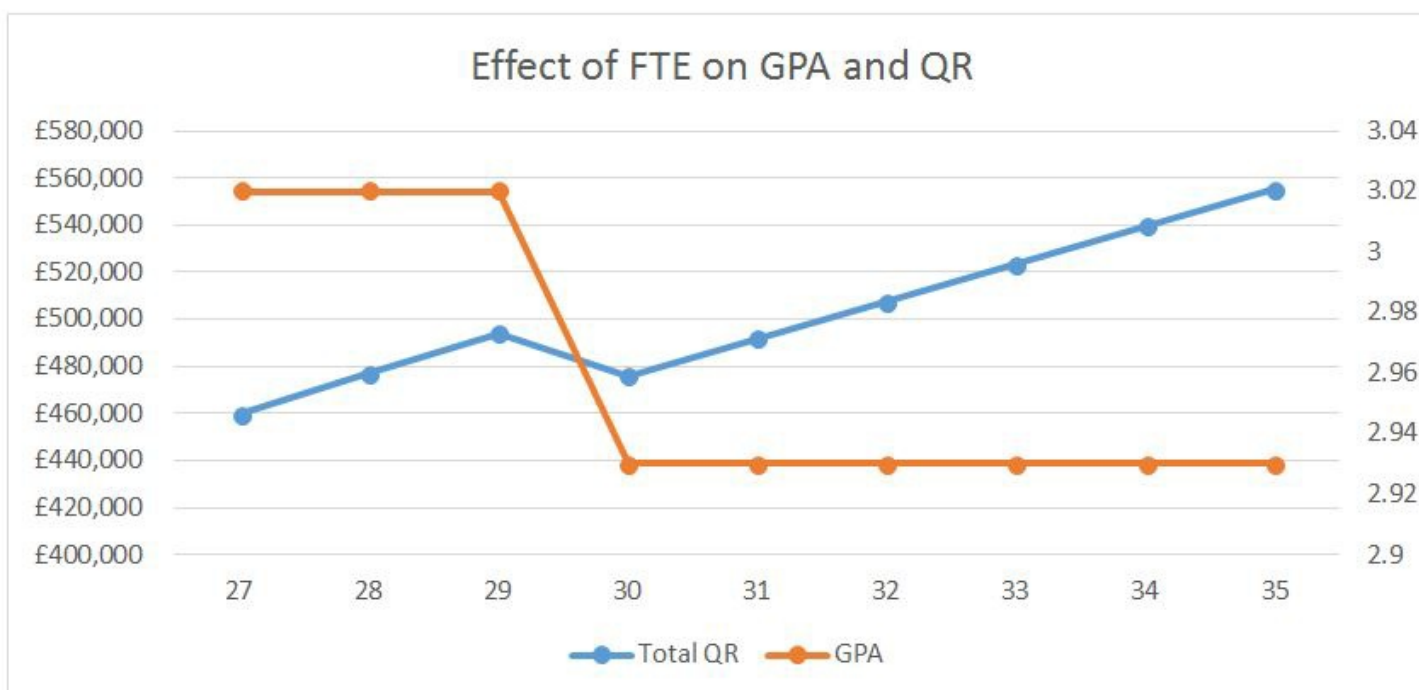
However, under the new requirements for impact case studies (one per submission “[plus one further case study per up to 15 FTE staff returned](#)”) the relative value of each case study (for submissions of 25 FTE or more) also goes up, even before the increased weighting of the impact component is factored in. On average this compound effect would be 70.10% (for the same distribution of submission sizes as REF2014; it is, of course, likely that the “all staff” requirement will mean more submissions of 25+ FTE, making 70.10% actually a lower estimate). So, using this lower estimate, the relative worth of impact case studies to outputs for REF2021 works out at 5.48:1; or to put it another way, the outputs of more than two people. The methodology for these calculations is [outlined here](#). It should be noted that calculations are based on a number of assumptions about funding, including the same basic unit of resource. As total FTE are expected to increase with “all staff” being submitted, the FTE unit of resource will decrease (unless there is a significant increase in overall funding!) But the main thrust of the argument should remain consistent: that the relative value of impact case studies as compared to outputs looks set to increase.

Perhaps an example will better illustrate the point. Using figures for a mid-range, mid-size submission we can extrapolate the effect of the new REF rules on quality profile and hence QR funding. Again, there are, of course, a large number of assumptions here that all other factors will stay the same; the major one being the quality multiplier, with an award of 4* being worth four times that of 3* in terms of funding last time round.

As our example, let's take a 29-FTE submission and use the HEFCE QR values for History from 2017/18. Assuming quality profiles of 4*:3*:<=2* (for simplicity I'll group everything under 3* into the 2* category) for Outputs of 24%:43%:33% (and the same for Environment), the GPA of these elements is 2.91, and we'll keep that fixed throughout the scenarios described below. Let's say we also have at our disposal four impact case studies graded at 4*, 3*, 3*, and 2*. This means that for a submission of 29 FTE we will select the best three case studies and hence get an impact profile of 33%:67%:0% (GPA 3.33). This contributes to an overall GPA of 3.02 and a linear Research Power (GPA multiplied by the number of submitted FTE) score of 87.44.

But if this submission were to include an additional FTE and we assume the average output (and environment) profile stays the same – in fact the output profile would presumably drop – then the FTE reaches 30 and hence the fourth impact case study is needed. This means that the 2* case study must now also be submitted, and so the impact profile drops to 25%:50%:25% (a GPA of 3.00). Combining this with the same output and environment profiles means the overall GPA drops to 2.93, and whilst the research power rises to 87.95 (due to the increased FTE), because of the dilution of 4* the QR allocation would actually go down! In fact, in this scenario, even increasing to 31 FTE still leaves us with less QR than the 29-FTE submission. Going the other way, a 28-FTE submission will also result in more QR than submitting 30 FTE (see Table 1 and Figure 2).

FTE	27	28	29	30	31	32	33	
ICs	3	3	3	4	4	4	4	
QR-Outputs £	266496	276366	286236	296107	305977	315847	325717	3
QR-Impact £	126874	131573	136272	105729	109253	112777	116301	1
QR-Environment £	66624	69092	71559	74027	76494	78962	81429	
Total QR £	459994	477031	494068	475862	491724	507586	523448	5
GPA	3.02	3.02	3.02	2.93	2.93	2.93	2.93	
Power	81.41	84.42	87.44	87.95	90.88	93.82	96.75	

Table 1: Effect of FTE on QR, GPA and linear Research Power. [Click to enlarge.](#)Figure 2: Effect of FTE on QR and Grade Point Average. [Click to enlarge.](#)

So while research power generally increases with FTE (though don't be fooled, for simplicity I've only used whole-number FTEs here; 29.9 FTE, for example, gives a research power of 90.15 – higher than for the 30-FTE submission), the GPA is badly affected, and even the QR takes a while to recover. To put this in context, in REF2014's History UoA a drop of GPA from 3.02 (30th) to 2.93 (38th) meant a drop of eight places – simply by adding one weaker case study. An even worse scenario would be a missing or unclassified fourth case study, which would drop the impact profile GPA to 2.50 and the overall to 2.82 (=49th). It would be a brave institution indeed that would be willing to risk that. The loss of £18k per annum – the difference between the 29-FTE and 30-FTE submissions in the example above – seems to pale into insignificance.

So, this means the driver for excluding staff to stay below an impact case study threshold will be even higher than last time. Previously, non-submitted staff had to carry the stigma of not having been selected. This time, the “mechanism” for non-submission will likely have to be a contractual change to teaching-only status; surely far more damaging?

Will institutions rise above these issues? Or will post-submission analysis again reveal a distinct lack of submissions with FTE sizes at the thresholds (e.g. 30, 45, 60, etc.)? History, if you will excuse another pun, suggests it will be the latter.

Note: This article represents the views of the author, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our [comments policy](#) if you have any concerns on posting a comment below.

Featured image credit: [Chess board](#) by Luiz Hanfilaque, via Unsplash (licensed under a [CC0 1.0](#) license).

About the author

Simon Kerridge is Director of [Research Services](#) at the University of Kent. He is the immediate past Chair of [ARMA](#), and the current Chair of [CASRAI](#), and is writing up the [RAAAP](#) project looking at the profession of research management and administration around the world. You can find him on twitter [@SimonRKerridge](#).