

## Application and Consistency of Approach in the Use of Performance Metrics

A report by the Associate Provost [Institutional Affairs]

December 2015

### 1 Introduction

1.1 In their review of performance management policies at the College, which was presented at Provost's Board in February 2015, the Director of HR and the (then) Senior Consul noted that:

*"... a number of concerns were raised ... about the application and consistency of approach in the use of performance metrics in academia and in the College."*

As a result, the Provost asked the Associate Provost [Institutional Affairs] to convene a small team to undertake a review of the application of performance metrics for academic staff at Imperial College, the recommendations to be submitted for consideration by Provost's Board. It was subsequently agreed by the Provost to restrict this review to academic staff (Lecturers, Senior Lecturers, Assistant Professors, Associate Professors, Readers and Professors) and to consider other academic researchers (PDRAs, Research Fellows and perhaps others), who are of major importance to the College but who nevertheless have their own (and different) concerns, at a later date.

### 2 Review Team

2.1 The review was chaired by Professor Stephen Richardson (Associate Provost [Institutional Affairs]). He invited the following to join the review team, all of whom kindly agreed to serve and all of whom provided extremely valuable input to our discussions:

- Professor Jan Cilliers (FoE),
- Professor Stephen Curry (FoNS),
- Professor Jonathan Haskel (ICBS),
- Professor Desmond Johnston (FoM),
- Ms Ann Kelly (HR),
- Mr James McSean (Provost's Office).

### 3 Current Practices in Faculties

3.1 Performance metrics can be and to some degree are used in the College to inform decisions about academic staff regarding their:

- appointment,
- probation,
- promotion,
- PRDP objectives,
- pay relativity,
- performance review and management.

3.2 Faculty Deans and HoDs (where there is not a single faculty practice) were asked to state their current practices with regard to performance metrics on any of these matters. ICBS and FoM each has its own faculty practice; FoNS is working towards one and FoE has different practices in each of its departments. The very wide variety of the responses is notable, ranging from a complete absence of metrics in a few cases to a very detailed and highly prescribed set of metrics in others.

## 4 Views of Academic Staff

4.1 The views of all academic staff were sought during September 2015 using the College's Qualtrics survey system, members of staff thereby being assured that all responses would be completely anonymous and confidential. Rather than asking specific questions, we asked one general question on performance metrics at Imperial. A comment box was provided for staff to share views on the application of performance metrics at the College. Comment could include but were in no way limited to feedback on their experiences of performance metrics and views on how they are applied, as well as any further thoughts or suggestions they would like us to consider.

4.2 Out of 1338 academic staff surveyed at the College, 309 (23% of the total) responded (another 217 started the survey but did not submit anything). All members of the review team read all of the comments made in the responses and we are most grateful to those who responded and greatly value their views. In summary:

- the response rate from each Faculty was approximately the same (ie the number who responded in each Faculty was roughly the same percentage of the total academic staff in that Faculty),
- about 15% of those who responded were Lecturers or Assistant Professors, 37% were Senior Lecturers, Associate Professors or Readers and 48% were Professors (note that, out of all academic staff at the College, 34% are Lecturers or Assistant Professors, 30% are Senior Lecturers, Associate Professors or Readers and 36% are Professors) - note that responses could not be broken down by gender or ethnicity, to ensure that no respondent could be identified,
- of those who responded, about 22% made comments generally in favour of the use of metrics, 33% made comments that were generally neutral and 45% made comments that were generally against the use of metrics (but note that these categories are rather broadly defined).

4.3 Short extracts from comments in favour of the use of metrics are given in **Annexe 1**: similar comments were made by many other respondents. Short extracts from comments against the use of metrics are given in **Annexe 2**: again, similar comments were made by many other respondents. No extracts from comments that were generally neutral are given - but it should be noted that many of these nevertheless expressed some reservations about the use of metrics

## 5 Performance Profile

5.1 In seeking to form a view on performance metrics, we started from the premise that, whatever their benefits or deficiencies, performance metrics pervade UK universities. From REF to NSS via the THE and their attendant league tables, universities are measured and ranked in many dimensions and any view of performance metrics has to be formed in this context.

5.2 The benefits and deficiencies of performance metrics are reflected in the wide variety of views in our survey of academic staff. Some members of staff would welcome them, because a well-elaborated set of measures is something clear that they can aim for: X papers in journals, Y raised in grants, Z scored in teaching surveys. Such concrete measures may also remove what some staff perceive to be arbitrary decisions on probation and promotion. That said, many of those generally in favour of the use of metrics also emphasised that they should be used with care, as part of a broader conversation about performance, and that their role in assessment should be clearly communicated to staff.

5.3 Equally, such metrics have some well-known deficiencies:

- true performance in a job can be hard to measure and can be gamed (what is a good paper? what if the teacher sets an easy exam?), so formulating targets purely in terms of metrics risks being vague or inappropriate or spuriously precise and risks distorting behaviour and undermining collegiality,

- academic work is multi-dimensional, involving research, teaching, administration and general citizenship: devising metrics that encompass all these dimensions is probably impossible, and runs the real risk of providing an incentive to neglect parts of the jobs that are unmeasured (eg collegiality and citizenship) and such desirable characteristics as creativity, leadership and risk-taking (eg starting a new research area),
- even if a set of measures and their levels could be devised, determining their relative weights would be very hard (and pointless since they are completely arbitrary) and would be likely to distort performance (eg if we over-value research funding, some scholars might move out of a field with little intrinsic chance of attracting funding).

5.4 Given this context, how are we to proceed without relapsing into the banal (let's all be very good at everything)? We have to acknowledge that, even if Imperial were to abandon internal performance measures, external measures (such as REF and NSS) would still continue. Whether we like it or not, such measures are inevitably used for financial allocation, influencing student preferences and affecting donors. To many outsiders, the College's reputation depends upon them. We cannot act as if they do not exist - and their very existence has inevitably shaped our strategy. In its deliberations in formulating its 2015-2020 strategy, the College decided that it wanted to be a world-leader in research, teaching and interactions with business and government. If we have that desire, we have to accept that, in a world of metrics, we will be held to account on the basis of demonstrable outcomes that students, sponsors and donors consider to be important (eg being in the top 5 in league tables, being in the top third in NSS, having over 30% of our income as commercial donations). Nevertheless, it is vital that we strive for real quality, not just a place in rankings. We need to have the courage and determination to define our own goals and to celebrate when we achieve them. If we get that right, performance as judged by metrics used by outside bodies should be recognised as a matter of course. To be truly world-leading, we ourselves need to be able to define and recognise quality.

5.5 We wish, as an institution, to excel in all of the range of external metrics upon which we are assessed. By extension, each faculty and each department must excel in all of them. We do not, however, feel that this can be taken to the individual level. We do not want to distort individual performance adversely - and it is probably unreasonable to expect individuals to excel at every single metric. This suggests that an approach to individual performance should be based on what we might call a *profile*, ie individual academics' performance would form part of a portrait of all of their contributions to the College. There would be no weighting of these contributions into one summary metric. Departments would need to have their own desirable performance dimensions to suit their subjects and circumstances (which would, for example, probably be very different for experimentalists and theoreticians). Individual colleagues would not be expected to excel in every dimension so measured - but there would be the firm expectation that they would make substantial contributions in all the main areas, namely teaching, research and administration (interpreted in a rather broad sense: see below).

## 6 Possible Profile Dimensions

6.1 What might the dimensions be that contribute to an individual member of staff's profile? In no particular order, they might include (but are in no way limited to):

- teaching: quality (as indicated by surveys eg SOLE and student nominations but also with a citation by the DUGS or equivalent to cover such matters as commitment and innovation),
- research publications: in journals, at conferences (refereed or not), in books (or chapters) etc,
- research income: in the context of the individual member of staff's research field (staff must aim to secure funding for their research – but precise amounts are much less important),
- quality of supervision and mentorship of doctoral (PhD & MD) students and PDRAs,
- national and international esteem: including but not limited to: election to top institutions (eg

RS, RAEng), conference keynotes/plenaries/session chairs (etc), active membership of editorial boards, medals/prizes, membership of government/national/international committees, active membership of learned institutions,

- administration: including being a good citizen (ie doing the jobs that make the College work - like being an HoD, DUGS, College Tutor or mentor; see also Imperial Expectations and Respect for Others), contributing to outreach etc,
- creativity, leadership and risk-taking (in order not to stifle creativity: metrics do not give a full picture of relative contribution to collaboration and multi-academic work; staff ought to be able to switch to new research areas, often with at least a short-term reduction in productivity, however measured),
- opinions of referees: national and international.

In all cases, the emphasis must be on quality before quantity.

6.2 In order to be effective, these dimensions (which are almost certain not to be uniform across the College or even within a faculty or department) would have to be published at least at departmental level (so that all members of staff know how they individually are being judged) and should be agreed with each member of staff on an individual basis: there should be no surprises for individual members of staff or for their managers.

6.3 How should those dimensions be quantified (if at all)? Once again, context is important. Academics are already measured by REF and by student surveys like SOLE - and such measures themselves comprise a range of sub-measures (eg publications and impact, lecturing and feedback). Thus, in developing a portfolio of measures to build a profile, it would be up to individual Departments to complement these existing metrics using judgement as they see fit. There is no one-size-fits-all and metrics would always have to be adapted to personal circumstances (such as part-time working). Indeed, we would expect the measures in the profile (many of which will be non-quantified and often non-quantifiable) to be adapted to fit each individual member of staff, perhaps as part of the PRDP process.

6.4 We would note that, whatever metrics are adopted as part of the profile of an individual member of academic staff:

- metrics can and must never be a substitute for proper (thinking and caring) management,
- one size does not fit all: there will be some metrics that are more appropriate for some members of staff than for others,
- the concept of responsible metrics (developed in the HEFCE Review of the Role of Metrics in Research Assessment & Management) must lie at the heart of any policy, as follows:
  - robustness: basing metrics on the best possible data in terms of accuracy and scope,
  - humility: recognising that quantitative evaluation should support - but not supplant - qualitative, expert assessment,
  - transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results,
  - diversity: accounting for variation by field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system,
  - reflexivity: recognising and anticipating the systemic and potential effects of indicators, and updating them in response.

## **7 Recommendations**

7.1 We want everyone at Imperial to contribute to the College's excellence and for the College to recognise every contribution. Such contributions should be the source of pride for both the individual member of academic staff and the College as a whole. Excellence is a multi-dimensional concept: some aspects, but by no means all, can be measured. Having some metrics could be helpful

to some staff members if they were in the form of clear and verifiable measures that might be subject to less uncertainty than academic judgement - but they would not be one-size-fits-all. Equally, having unclear or incomplete metrics would send out the wrong signals, driving all the wrong sorts of behaviour. Moreover, it may be unreasonable, other than in truly exceptional cases, to expect each individual member of staff to excel in every single dimension of their performance (though we need to guard against selfish opportunism by prima donnas).

**7.2 We therefore propose and ask Provost's Board to agree that:**

- **each department should develop profiles of its academic staff based on a series of published (ie open and transparent [perhaps on the College intranet]: there must be no surprises) dimensions, some of which can be measured and some (perhaps most) of which will quite properly rely on peer judgement:**
  - **each individual member of academic staff is not expected to contribute at an equal level to every single one of the many dimensions of such a profile - but is expected to contribute to all of the key ones, associated with teaching, research and administration,**
  - **the relationship of each individual member of staff with respect to these profiles be agreed by the HoD or a nominated substitute with each member of staff,**
- **building on the proposed profiles from each department, the Dean in each faculty (with assistance from HR) should agree the profiles with each department by the end of June 2016, ready for implementation at the start of academic year 2016-17,**
- **the process should be reviewed by the faculties, reporting to Provost's Board by June 2017,**
- **any resulting changes to the process should be ready for implementation at the start of academic year 2017-18.**

## **Annexe 1 - comments in favour of the use of metrics**

- *Performance metrics are key to ensuring the highest standards of the College. They should give priority to the production of innovative and impactful research.*
- *Performance metrics are helpful and should provide guidance to the individual as to the general standards expected. However they should not be used to make final decisions.*
- *Performance metrics, if used intelligently, can help people to develop their careers.*
- *In my department we use metrics in a considered manner which takes into account the whole of the academic activity: teaching, research and service. It is imperative that we maintain the ability to balance these.*
- *Clarification of what would lead to a top score in service, at every level, is really needed.*
- *Metrics are useful and good when the staff being evaluated and evaluators are aware of the subjectivity of metrics and when the staff being evaluated have the necessary support to improve his/her metrics.*
- *To identify a specific metric for each activity will be challenging. However, explaining what activities are of particular importance to the College appears useful.*
- *The promotion process is very opaque in my department and any way that this can be made more transparent would be an improvement.*
- *As a HoD I use metrics although one size does not fit all, even within a department. I think it important that one sets expectations - however it leads to people skewing their work sometimes in a way that is not optimal.*
- *Performance metrics need to be part of a continuous conversation at faculty and department level.*
- *Performance metrics can be a useful source of quantitative information but should be treated with great care.*
- *It's appropriate for performance metrics to be discussed, for example in PRDPs, alongside other information such as administrative roles, teaching and outreach but they should not be the only consideration and they should not be used without careful referencing to the norms for that subject area.*
- *It is important to have clear guidance on the criteria for any metrics and how they are to be applied fairly and consistently across faculties.*
- *A fair assessment of individual academics requires a very much case-by-case and field-by-field application of metrics, where qualitative arguments should play a role as well.*
- *Metrics are used at the point when a member of staff is significantly underperforming to pressurise that individual to either radically improve or leave. How to improve this? Make clear what level of performance is required at different levels. Make those measures relevant to researchers and flexible. Train managers in people management. Set up a mentoring scheme for all staff who would like a mentor to provide support for career development, work/life balance.*

## **Annexe 2 - comments against the use of metrics**

- *All my colleagues at MIT and Harvard etc tell me they reject metrics because they lead to mediocre candidates. If Imperial really wants to be a leader, it has to be bold enough to judge based on quality.*
- *Comparison between different groups in a department suffers from the normalisation problem and across subjects even more so.*
- *There is an impression that performance metrics are a management tool and there is a risk that the true value of scholarship will be eroded or missed.*
- *Any rigid system tends to distort behaviour.*
- *The performance metric mentality is generally liked by the central services of institutions/ companies as the data are readily available. It certainly keeps individuals distant from the decision makers and it can lead to misinterpretation of people's contribution.*
- *College should minimise use of metrics at the individual level. They should be regarded as weak indicators of performance by themselves and more holistic approaches should be encouraged.*
- *Academics have a wide range of styles of achieving and in my view it is best to judge these in the round, bearing in mind their subject area and personal strengths and weaknesses.*
- *One has to be very careful to make clear that performance metrics are used as a guide in appraising performance and contribution, not as a means of making decisions.*
- *Metrics should promote collaboration not competition: this is a major strength of Imperial and there is a risk that metrics could undermine it.*
- *Performance criteria based on funding success, especially over a short period, are unreliable because obtaining founding nowadays is such a lottery.*
- *Needs to be a mix of formal targets and informal academic judgement on performance.*
- *Use of standard metrics is generally a substitute for good performance management. Good management is about being present, understanding your staff and how they work together.*
- *Primary performance metrics seem to be currently narrowly focussed on publishing in prestigious venues and obtaining research funding, ideally as PI. Metrics encourage neglect of teaching duties, a general disengagement of academics from students and any activity that is not metric-related.*

- *Metrics should not be designed to discourage risk-taking: we need to be bolder, more creative and take more risks as an institution.*
- *A fair assessment of individual academics requires a very much case-by-case and field-by-field application of metrics, where qualitative arguments should play a role as well.*
- *Creating performance metrics that occupy the time and attention of all staff, just to help deal with a small number of under-achievers, is collective punishment.*
- *Use of metrics across disciplines can lack interpretability and lead to dangerous and unjustified conclusions.*
- *There is definitely a culture creeping in of only doing what counts that makes it very hard to get staff to participate in necessary but ad hoc activities*
- *I have only ever seen performance metrics used for one purpose - which is to provide an excuse to fire staff.*
- *I am line manager for several junior staff going through probation and I think the performance metric culture is poisonous for them. That is unfair for probationers who see others performing far worse than them but without censure, while they are held to a higher standard.*
- *Performance metrics attempt to measure the performance of academic staff in quantitative terms. In my opinion, this is very problematic, given the nature of the tasks we perform.*
- *Over-reliance on them will lead to metrics being deliberately gamed (not difficult).*
- *The search for one-size-fits-all metrics is the worst thing about Imperial and frankly makes us the laughing stock of other universities. Managers should be competent to make judgements about performance and not rely on the tick box mentality that Imperial likes. In my experience it is applied in a heavy handed, often bullying way.*
- *Assessing performance is reasonable. The use of metrics as a short cut is indefensible and undermining the culture. The whole process is corrupting science.*
- *Not only inhumane but also inefficient and destructive. It focuses too much on money as well as impact factors but not on the quality of the science itself.*
- *Quantitative characteristics are not good for measuring qualitative things.*
- *To allow academic the freedom to engage in truly cutting-edge work they have to be allowed the possibility of failing for extended periods of time.*
- *Performance metrics fail to capture long-term intangible benefits and lead to increased pressure and short-sighted goal-oriented targets.*
- *Performance metrics save top management from consulting with peers and from having to spend time and effort understanding specific fields and contributions. This is why top places rely on peer reviews instead.*
- *As a faculty, we are by definition a group of peers who trust each other's judgement to work together, therefore we shouldn't be relying on performance metrics to grade one another, rather peer-assessment from our colleagues.*
- *Metrics never represent the context in which they have been measured, they do not represent balance and role in a complex system such as academics in a university.*
- *In my long experience, the most useful performance indicators are actually quite hard to capture numerically and from simple data. For example, leadership, originality, vision, ability to generate, organise and drive (in research, teaching, professional services, administration).*
- *Performance metrics encourage the short-term pursuit of funding and quantity of publications rather than the longer-term pursuit of knowledge leading to scientific breakthroughs.*
- *Rather than quantify and micromanage the performance of academics, why not try to produce a nurturing, exciting and supportive environment which encourages people to develop new ideas? It seems that discredited management approaches are being applied which stifle creativity and true innovation.*
- *The tagging of metrics to individuals fails to recognise that research and teaching is delivered by teams who at varying times will have varying roles in each of these domains.*
- *The use of performance metrics is no substitute for responsible and intelligent management, and sometimes directly undermines the process of good management.*
- *Performance metrics are partly to blame for an increasingly self-centred breed of academics, and the cause of a wider erosion of collegiality.*
- *How do you balance and account for individuality and the ultimately individual nature of research, balancing short-term goals and long-term effects, effectively tackling the not easily measurable, defining priorities in a sufficiently broad-based and flexible way, ensure the alignment of individual and institutional goals and maintain a happy and exciting environment?*