Leader in Higher Education Policy Research
Policy Research | Value Creation | Innovation
www.ipptn.usm.my
UNIVERSITI SAINS MALAYSIA
APEX
CENTRE OF EXCELLENCE | UNIVERSITI SAINS MALAYSIA
IPPTN
INSTITUT PENYELIDIKAN PENDIDIKAN TINGGI NEGARA
NATIONAL HIGHER EDUCATION RESEARCH INSTITUTE
Institut Penyelidikan Pendidikan Tinggi Negara
ISSUES PAPER
National Higher Education Research Institute
No. 05 | 2021

Quality, Excellence, and Impact: Can We Really Measure Them?

Quality, Excellence, and Impact are the holy trinity of higher education. They are the holy grail that every university and academic is striving towards. In some instances, they are together and referred to interchangeably, while in other instances, one or two of them may be mentioned independently. Collectively, they are almost omnipresent in policies, regulations, mission statements, written or verbal statements about universities, as well as used as the benchmark for what the university does in terms of developing, disseminating, and applying new knowledge.

Before we proceed, it is imperative to define the context for which we shall be discussing quality, impact, and excellence. First and foremost, a university is an institution that must be rooted in the society to which she has been entrusted to preserve, understand, advance, and disseminate intellectual, scientific, and artistic knowledge across generations. This societal mandate is then translated into the usual functions and activities of an academic that evolves around research, teaching, and services.

The Concepts

The notion of **quality** in education has been a long-standing debate. Harvey and Green (1993) argued that "quality is not a different perspective on the same thing, **but different perspectives on different things with the same label"** (p. 10). In other words, quality means different things to different people, and equally important to recognise, the concept of quality can also be relative as well as absolute. For instance, when a university claims to provide quality education, the term quality can be understood differently by the management of the university, academics who teach, students who will experience the education process in person, parents who pay for the education, bureaucrats and regulators who regulate the programme, as well as employers who are expected to hire these students when they complete the educational process. Furthermore, to some of those mentioned above, quality is a comparative or relative measure, while to others, a precise and specific measure.

While the concept of quality is more commonly associated to teaching and learning activities, the concept of impact mostly refers to research activities. Yet, likewise, **impact** also suffers from the same conceptual problem as quality. Impact can be formally defined, such as how the Research Excellence Framework (REF) in the United Kingdom (UK) has done in 2014, whereby impact refers to "an effect on, change or benefit to the economy, society, culture, public policy/services, health, the environment or quality of life, beyond academia" (UKRI, n.d.).

However, the concept of impact has not shown directionality and causality on the effect, change or benefit deriving from research activities, as if whether the effect/change is positive or negative and who benefits as well as affected by the effect/change.

Importantly, the concept of impact, as defined by UK Research and Innovation (UKRI) for REF as well as commonly used in the discourse of higher education, has almost exclusively centred on research (more specifically on knowledge advancement), and has not been used to refer to other equally crucial roles of a university on knowledge preservation, understanding, and dissemination. Like quality, impact also has the tendency of being used as the same label from different perspectives to refer to different things.

Excellence is a concept that is traditionally intertwined with credibility in the academic profession, that holistically includes the functions expected of the university and individual academics (Becher, 1989; Morley, 2002). Credibility reflects academic excellence as the highest accord that is defined as appreciation, respect, and acknowledgement given by fellow experts who are independent (Azman and Abdul Rahman, 2017). In other words, academic excellence and credibility are concepts built on

essential traditions of academic profession: peer (expert) review and collegiality.

Metrics and Measurement

Traditionally, the articulation of quality and excellence of a university, and to some extent impact, was left very much to the academic community to evaluate these concepts collegially through review by fellow experts. However, as universities are now seen as 'knowledge producers' and agents for economic growth, and coupled with the shift in many developed and developing countries to channel public monies to support a range of activities in higher education, therefore the articulation of these three concepts became much more important and needed to be explicit.

Concurrently, this shift has also been motivated by the infiltration of the neoliberal ideologies that emphasised accountability through metrics and measurables. Hence, in the last two to three decades, quantification has become the norm in almost every aspect of higher education, and the term Key Performance Indicators (KPIs) has become the buzzword in the governance, management, and administration of universities.

The launch of global university rankings – Academic Ranking of World Universities (ARWU) (more commonly known as Shanghai Jiaotong Ranking (2003), Webometrics (2004) and Times Higher Education World University Rankings (2004) – was also a watershed and catalytic moment that sparked the shift into using standardised metrics and measurements as indicators of quality, excellence, and impact.

Although quantification of indicators and the use of metrics have the potential to strengthen governance, management, and administration by turning these accountability measures to be more objective, there are many fundamental issues and myths about metrics and measurement that needed unpicking.

Measurement Myths on Quality

As Ashwin (2020) argued, there are three measurement myths about quality in the context of higher education. First, university rankings, while trying to appear as transparent, purposeful to measure quality, objective, and rigorous; are in fact the opposite. These measures are meant to sell things for commercial purposes such as advertisement and consultancy services, favour the higher status institutions, and are heavily based on reputational survey from selected groups of academics and employers.

Second, in the era of big data, there is belief that the measurement of students' experience, particularly learning analytics, can be combined to illustrate quality. Yet, many of the indicators, metrics and measurements have been taken out of the educational context, as often we assess what is measurable, and we rely on the data that are available rather than the data that importantly reflect quality.

Third, there is a danger that once a measure becomes a performance indicator (outcome), it ceases to be a good measure. This is a common problem because those who are being assessed, regardless of individuals or institutions, will change their practices in order to maximise their performance. Inevitably, we will have a situation where performance for that specific measure rather than performance for what the measure is intended to capture.

For example, if the number of contact hours lecturers have with their students is the yardstick of quality, the rational approach to enhance quality therefore will be to increase the number of contact hours. But will quality of learning for students increase if lecturers allocate more time for students to meet them? Another more pertinent example would be number of publications, whereby it is possible to boost the number of publications without actually trying to produce publications that are useful. In other words, the measurement has become something that can be gamed, instead of genuinely reflect quality.

Barthes (1957) aptly summarised the myth about measurement/quantification on quality, whereby he argues that the myth simplifies reality and this myth only measures reality by numbers, not by quality.

The discourse about quality in higher education needs to further extend into the aspect of quality assurance. Quality assurance, in the recent decades or so, has become a major feature in higher education, where regulatory agencies, such as the Quality Assurance Agency for Higher Education (QAAHE) in the UK or the Malaysian Qualifications Agency were established to regulate and assure quality. However, the elephant in the room that is rarely asked, "What exactly is quality and how quality is measured?"

The QAAHE in the latest Quality Code defined 'high quality' as "quality which can consistently lead to credible and recognised positive outcomes for students. High quality is the minimum level of quality that is expected of all providers of UK HE" (QAAHE, 2018, p.5). What a wordy and somehow confusing definition, but the keywords that required our attention are: **outcomes and minimum level**.

Outcomes and level descriptors have been used to express and operationalise quality in higher education, influenced by the managerial principles of predictability, consistency, and standardisation, as well as believed to be more transparent, focused,

comparable, and importantly more accountable for public monies used in education (Avis, 2000; Jackson, 2000; Scott, 1995; Wolf, 1995). However, outcomes disassociate the processes of learning, where only the 'what' is mention and the 'how' or 'why' is ignored.

Initially developed in the United States in the 1960s as competence-based (job-specific) assessment in the area of teachers' education, outcomes, like square pegs being forced in a round hole, was adapted into higher education and used as a generic one-size-fits-all instrument, including in non-competence-based programmes and courses under the pretext of quality assurance (Wolf, 1995). Knowledge acquisition and understanding, developed through reading subjects like history and philosophy, are among the ill-fitted domains that are now required to be stated in terms of outcomes.

Furthermore, even usage of outcomes in competence-based education has its limitation whereby 'standards' and levels of proficiency are only prescribing the minimum pre-requisite in a pass/fail dichotomy to the extent minimum pre-requisite will become the maximum standard to achieve (Wolf, 1995).

Outcomes, certainly, do not promote mastery of the competences, neither do they address a fundamental dimension of learning and education which is behavioural as Burke (1995) quoting Stenhouse that "education as induction into knowledge is successful to the extent that it makes the behavioural outcomes of the students unpredictable" (p.60).

Conversely, apart from the attempt to associate quality with outcomes, another direction in which the notion of quality has 'deviated' is to use it interchangeably with accreditation. For example, the Code of Practice for Institutional Audit (COPIA) (MQA, 2008a) and Code of Practice for Programme Accreditation (COPPA) (MQA, 2008b) have explicitly articulated that quality in higher education in Malaysia is approached through accreditation of programmes and qualifications, and audit. What accreditation entails is the stating of the minimum criteria and standards for programmes and qualifications, and the audit is to then ensure the minimum criteria and standards that have been spelt out is fulfilled.

For example, a Bachelor of Science programme has 130 credits and several declared Programme Learning Outcomes, or a course in the programme has listed several Course Learning Outcomes with a list of topics and matters of assessment. Quality is then assumed to be assured when the Outcomes that were stated have been declared as achieved. Is that quality of experiencing an educational experience?

Measurement Myths on Excellence

The Research Excellence Framework (REF) in the United Kingdom (UK) is a well-known instrument to evaluate research of the British universities that was introduced in 2014 to replace the Research Assessment Exercise (RAE). A year later, an independent review was conducted to look at potential uses and limitations of research metrics and indicators, which produced a report titled "The Metric Tide" (Wilsdon et al, 2015). One of the key aspects this review focused on was the changing ways in which universities are using quantitative indicators in research management systems that have been strongly powered by rankings, as well as importantly, negative and unintended effects metrics have on the research culture.

"The Metric Tide" was chosen as the title of the report of this independent review in the UK to reflect powerful currents that have been whipping up the metric tide. These currents include pressures for audit, demands by policymakers for information on quality and impact, advancement in analytical capability as well as competition for prestige and resources.

However, despite these powerful currents, the independent review argued that peer review, despite its flaws and limitations, remain a key aspect of excellence and cannot be replaced by metrics. Even after more than a decade, the REF continues to rely on 34 sub-panels of expert assessors who took almost a year to evaluate tens of thousands of research outputs for its 2021 exercise. It was reported that for REF 2014, the onerous task of peer review cost GBP 19 million worth of the time of these experts.

The attempt to substitute metrics with peer/expert review has been driven largely by development of indicators for rankings purposes (Grove, 2021; Usher, 2017; Wilsdon et al, 2015). It is important to recognise that major rankings such as QS and Times are owned by corporations who created these metrics to determine and control what counts as quality and excellence in higher education. We need to recognise that this subtle but fundamental shift is happening whereby commercial entities like Quacquarelli Symonds (QS), Clarivate, Reuters, Elsevier and Times have replaced the peer/expert colleagues in deciding what is excellence.

If trading companies like the East India Company and Dutch East India Company played a major role to colonise and expand their empires in the 17th, 18th, and 19th centuries, are universities on the verge or has been colonised by these new forms of commercial entities? (see Wan, 2021).

Measurement Myths on Impact

It is important to note that REF is an instrument designed not only to evaluate research, but more specifically, research impact. It is equally crucial to note that although REF is the instrument to evaluate research impact, impact in the REF exercise is evaluated in the form of **narrative case studies**, divided into cultural, environmental, legal, societal, and technological.

In the 2014 exercise of REF, a total of 6,975 impact case studies were submitted to show the impact of each university in areas ranging from research to society at large. As Wilsdon and colleagues (2015) clearly pointed out, "for the impact component of the REF, it is not feasible to use quantitative indicators in place of narrative impact case studies, or the impact template" (p. x). They further cautioned that if impact is to be measured through quantitative indicators, such attempt is likely to constrain the thinking around which impact stories have greatest currency and constraining the diversity of research in the UK. In other words, attempt to measure impact will be counterproductive.

Another measurement myth surrounding the concept of impact is the indicators such as citations, journal impact factors (JIFs) and h-index. The h-index was developed as recent as 2005 by a physicist to measure relative quality of researchers through a single number that claim to measure productivity/ quantity and visibility (Bornmann & Daniel, 2007).

Likewise, the JIFs is an index that is based on two elements: number of citations in the current year to articles published in the previous two years, divided by the number of substantive articles published in the same two years (Garfield, 2006). JIFs of journals in the same field are then ranked and categorised into four quantiles to then classify the journal as Q1, Q2, Q3 or Q4.

It is important to note that while these impact measures, be it h-index for individuals or JIFs for journals, are essentially measuring citation of the papers published in indexed journals. Citation, in fact, is just referring to the number of times a paper is mentioned. Thus, it is simply a measure of visibility (not necessarily positive) and misleadingly used as a proxy for impact.

As such Wilsdon and colleagues (2015), in response to the use of indicators in REF, have questioned the robustness of indicators such as JIFs and considered them to be downright dubious. Understandably, therefore, the impact component of REF has maintained the use of narrative case study to illustrate impact and resisted any use of measurable metrics, because impact just cannot be measured. To put it bluntly, the h-index of an individual academic is the same measure as the number of likes someone has on Facebook. What does that got to do with impact?

Universities and Academics Must Stop "Doublethinking"

George Orwell first introduced the concept of 'doublethink', which refers to **the power of holding two contradictory beliefs in one's mind simultaneously and accepting both of them**. In the novel Nineteen Eighty-Four, doublethink was described as "to tell deliberate lies while genuinely believing in them, to forget any fact that has become inconvenient and then, when it becomes necessary again, to draw it back from oblivion for just so long as it is needed, to deny the existence of objective reality and all the while take account of the objective reality which one denies" (Orwell, 1949, p. 171). Orwell envisioned in his novel a society controlled by doublethinking, where it is maintained by an unconscious refusal to examine the assumptions and facts supporting one's beliefs, resulting in contradictory slogans like "War is Peace", "Ignorance is strength" or "Pro-child/Pro-choice" (Boss, 1993).

Ashwin (2020) aptly pointed out how universities and academics have been doublethinking in respect to the whole fiasco of university rankings. While universities and academics know rankings are nonsense, yet they still 'celebrate' when there is success for them. To some institutions, 'successes' in rankings are consciously or unconsciously included into marketing and promotional activities. To a large majority of institutions, academic performance evaluation has been gradually or drastically modified to be in lined with the definition of quality, impact, and excellence provided by external commercial entities either explicitly or implicitly for ranking purposes.

Universities and academics, pathetically, have either been forcibly converted to become believers of the doctrine of quality, impact, and excellence defined by external parties and become crusaders to strive towards capturing the holy grail of higher education through these 'objective' measurable metrics as the absolute truth. Or conversely, they have taken a defeatist attitude that the holy grail can only be achieved in the mould and measurement determined by external parties, whereby university leaders blamed that there is nothing left to do but to conform, and academics can only either "conform or leave" by taking on the mantra of "publish or perish".

If the university as the societal institution entrusted to understand, advance, and disseminate the intellectual, scientific and artistic knowledge, which also housed perhaps the largest concentration of learned scholars and public intellectuals, continue to

practice and cordone doublethinking on its fundamental pillars of quality, impact, and excellence; how can we expect the university to pass on these intellectual, scientific and artistic heritage to the future generations?

Moving Forward

As the saying goes, admitting that we have a problem is the very first step before we can talk about fixing the problem. University leaders and academics need to admit that we have been doublethinking, debunk the myth and stop such paradoxical mindset, before we can clear our mind to understand and re-think what exactly quality, impact, and excellence in our intellectual endeavours are. Only with such a realisation that we can reset the compass and chart our way forward.

The reset of the compass requires the need to return to the fundamental principles of what a university and an academic is. An academic is one who is tasked to understand, advance, disseminate and preserve a specific field of knowledge, and he or she carries out research, teaching, supervision, and engagement activities. The university is then the institution that provides the platform for individual academics to carry out the above activities collectively. Thus, the thinking of quality, impact, and excellence, must be built on this understanding of an academic and a university.

First, it is important to recognise, **the academic traditions to determine quality, impact, and excellence have been based on the principles of peer review and collegiality.** For instance, the highest accolade in the scientific community is the Nobel Prize. The awarding of Nobel Prize is determined not by number of papers, amount of grants nor number of citations. It is essentially determined by a committee comprising of experts in the field who are mostly past winners, and the determining criterion is the excellence of a scientific idea. In most cases, the evaluation is based on a book, an article or a series of coherent publication that adequately illuminate the quality and impact of the scientific idea by a group of fellow experts. That is peer review and guided by a sense of professional collegiality whereby there should not be a sense of jealousy or rivalry.

Likewise, the peer review tradition in the production of academic work, be in books or articles in journal, is the common practice to govern quality and excellence. The sense of professional collegiality is also important, as reviews and comments from reviewers are meant to be constructive to improve the work, idea, and publication in ensuring quality and excellence.

Similarly, as most academics today have a Doctor of Philosophy, the main determinant to award the highest academic degree and to show that the person is qualified to be called a doctor, is an expert who is not involved in the development process of the candidate. It is not the supervisor nor the university that determine whether a PhD candidate passes or fail. It is the internal and external examiner and that is essentially peer review.

Thus, the articulation of quality and excellence of an academic must be premised on the evaluation and recognition by another independent expert, and not by external entities who may have vested or conflicting interests. The recognition of quality and excellence on every single academic will collectively build on the department, faculty, school and ultimately the university.

Importantly, the articulation and recognition of quality and excellence cannot be measured by any metrics because these measurements are meaningless. It is not the number of papers that show quality and excellence, it is the way in which a body of knowledge has been understood or advanced that matters. It is not the number of hours spent teaching in a lecture, it is the way students are transformed through the interaction with the body of knowledge through the lecturer as a facilitator of such an educational experience. In a nutshell, it is the substance that defines quality based on the fundamental functions of an academic and a university, and excellence that is recognised by other experts, that truly matters. All these cannot be measured by indicators and metrics!

Second, the university, ranging from its leaders and academics, **must have the courage and confidence in itself to remain steadfast to the *raison d'état* of its very existence**. The university must be absolutely clear of its role and purpose, why the institution was established and to whom it has an obligation. In the craze to chase rankings by publishing large number of academic papers in commercially controlled indexed journals to gain citation (which falsely represent impact), the more pertinent questions to critically ponder: Has the advancement and understanding of knowledge used to change and benefit the society in any way? Has the environment in the campus benefited from the advancement in technology developed by the university? Has the socio-economic standing of the neighbouring community improved due to the knowledge advanced by the university? Has the university been able to consolidate all its resources to produce diagnostic kits or vaccine in the ongoing COVID-19 pandemic? We need to recognise that the narrowly focused recognition of impact through citation, has only benefited the profit margin of the supra-national commercial entities.

As Wilsdon remarked in an interview on the report The Metric Tide, there has been too much emphasis placed on measuring impact and in turn, creating a lot of perverse and negative effects in the knowledge system by forcing researchers to a narrow publication path dictated by commercial entities. Hence, this choked up the system by reinforcing the 'power of high impact' journals at the expense of other ways of dissemination of knowledge, and in the longer term, will be corrosive to the diversity, vibrancy, and relevance of the knowledge system of a university.

Thus, the only way for the university to have real tangible impact, is not to measure and conform to metrics, but to steadfastly focus on advancing, understanding, preserving, disseminating and translating the knowledge to the society. Importantly, this confidence and courage also suggests a university not to succumb to the pressure of comparison, both internally and externally in trying to outdo what other institutions are doing, because comparison and measurement is a distraction to the university in sticking to its *raison d'état*, role, and purpose.

In conclusion, let us be reminded by the famous words of Albert Einstein, **"not everything that can be counted counts, and not everything that counts can be counted"**. If we begin to measure the quality, impact, and excellence of Albert Einstein using the current bunch of metrics and measurements, where he authored 147 articles between 1901 and 1955, have 1,564 citations (only 27 self-cited) with an h-index of 56 (Gingras and Khelfaoui, 2020), he may not even get hired by a university or promoted to professorship in our current system. Do we consider ourselves to be of better quality, more impactful and more excellent than Albert Einstein, just because we have more published papers and higher citation?

References

Azman, N. and Abdul Rahman, S. (2017) Academic *Profession and Culture (AL201)*. Bandar Enstek: Higher Education Leadership Academy.

Barthes, R. (1957) *Mythologies*, Paris: Editions du Seuil.

Becher, T. (1989) *Process and structure in higher education*. London: Routledge.

Bornmann, L. and Daniel, H.D. (2007) What do we know about the h Index? *Journal of the American Society for Information Science and Technology*, 58(9), 1381-1385.

Boss, J.A. (1993) Pro-Child/Pro-Choice: An Exercise in Doublethink? *Public Affairs Quarterly*, 7(2), 85-91.

Garfield, E. (2006) The history and meaning of the Journal Impact Factor. *Journal of American Medical Association,* 295(1), 90-93.

Gingras, Y & Khelfaoui, M. (2020) *Albert Einstein the mediocre: Why the h-index is a bogus measure of academic impact*. Accessed 27 April 2021 at https://phys.org/news/2020-07-albert-einstein-mediocre-h-index-bogus.html

Grove, J. (2021) Creator says REF should swap expert panels for metrics in science. *Times Higher Education*. Accessed 29 April 2021 at https://www.timeshighereducation.com/news/creator-says-ref-should-swap-expert-panels-metrics-science

Harvey, L. and Green, D. (1993) Defining quality. *Assessment and Evaluation in Higher Education*, 18(1), 9-34.

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I, Wouters, P., Hill, J. and Johnson, B. (2015) The Metric Tide: *Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. DOI:10.13140/RG.2.1.4929.1363

Morley, L. (2002) Subject to review: Engendering quality in higher education. *Journal of Education Policy*, 16(5), 465-478.

United Kingdom Research and Innovation (UKRI) (no date) *REF Impact*. Accessed on 30 March 2021 at https://re.ukri.org/research/ref-impact/

Usher, A. (2017) A short global history of rankings. In E. Hazelkorn (Ed.) *Global rankings and the geopolitics of higher education: Understanding the influence and impact of rankings on higher education, policy and society*. Oxford: Routledge, 23-53.

Wan, C.D. (2021) *Breaking Free of the Captive Mind. The ACU Review*. Accessed on 27 April 2021 at https://www.acu.ac.uk/the-acu-review/captive-minds/

Prepared by:
Wan Chang Da, IPPTN, USM