

Measurement changes what is being measured: What does this mean for the measurement of quality and productivity in Higher Education?

Alan Bain and Nick Drengenberg

Introduction

In this presentation we take up issues related to the measurement of productivity and quality in higher education by questioning whether it is currently possible to do so. We will present what we believe to be a provocative, albeit well supported account of the current status quo and offer up an alternative approach to the way standards are employed to measure learning and teaching quality and productivity in higher education. We describe how quality and productivity are fundamentally linked and cannot be addressed individually. The presentation will include examples of key claims and propositions.

Standards and Frameworks

When developing or working with a learning and teaching standards framework we are all hoping for or assuming those efforts will exert a positive effect on student outcomes. We expect this to occur as a result of a relationship among standards, practice, quality and productivity that ultimately influence the student experience in positive ways. In general terms we expect the relationship described simply in figure 1.

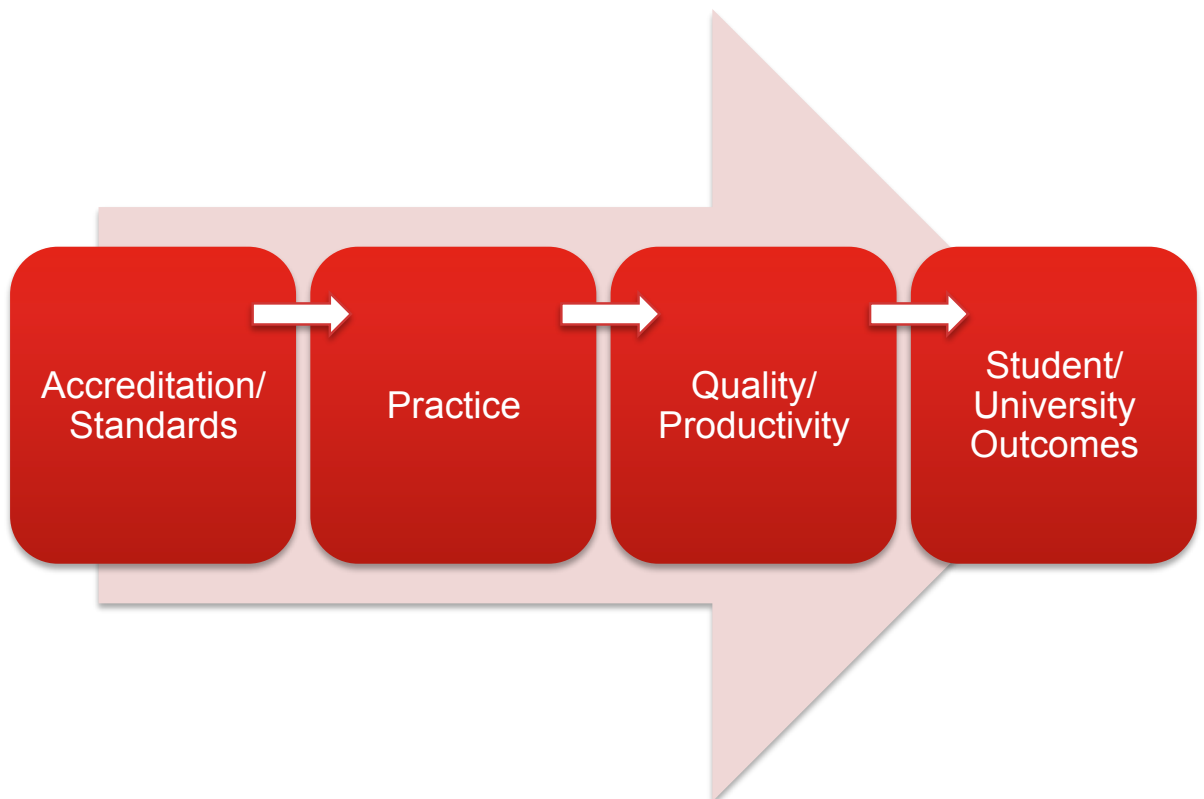


Figure 1: Expected Relationship driving a Standards Framework

In this presentation we ask and attempt to answer some serious questions about the nature of the connections illustrated in Figure 1. In particular, we take up the question of *what do we measure*, when establishing the connections across the elements and how well those measures line up with *what we need* in order to determine the quality and productivity of learning and teaching.

Questions

At the heart of our concerns are a series of questions about assumptions surrounding the elements in figure 1.

Question 1:

Do we understand the relationship between student learning and teaching practice at scale in higher education?

Such a relationship would require an understanding of what we do. According to Bowker & Starr (2000), this requires the existence of ***scalable visible and comparable professional practice*** that generates the kind of framework of professional control seen in fields like medicine and engineering etc. Such frameworks for practice when applied to learning and teaching could then be related to student outcomes.

There is no evidence to suggest such a foundation of practice currently exists in higher education. Contemporary research would suggest we know little about the relationship between learning and teaching. Most of the variance in students' performance in university is predicted by their characteristics at entry. There is much more variability in terms of student learning within universities than between them indicating the lack of variance contributed by a university's unique approach (ACT, 2009; Liu, 2008; Pascarella & Terenzini, 2005). Applied research investigating perceptions of learning and teaching in the area (e.g., Scott & Scott, 2015) would suggest that the idea of scalable learning and teaching practice in higher education is not a widely accepted idea even among learning and teaching leaders (e.g., Scott, Coates & Anderson, 2008). A novel corroborating perspective is represented in the following quote from (Mawdsley & Cumming, 2008).

"The lack of acceptable standards for pedagogy and classroom methodology is the key reason for the difficulties applying the concept of educational malpractice to ineffective teaching in litigation in the United States, the United Kingdom, and Australia."

The rejection of professional control is also not only a problem of management. Academic staff working in Higher Education institutions frequently retreat to the relationship with students to defend the status quo as well, often along the lines of respecting the individual learning characteristics of students, and arguing that any form of standardization and professional framework is an attack on professional autonomy. In our book *Transforming the Measurement of Learning and Teaching in Higher Education* we show how this is in itself a standard phase that other professions have moved through in their path towards the

development of a professional framework. We flatly reject the idea that a professional framework in any way compromises the ability to provide individualized professional ‘care’, and in fact argue the opposite – that it is impossible to specifically tailor the professional intervention required with what we will describe as the workable distinctions in quality provided by such a framework.

Question 2:

Do current standards really drive valid workable distinctions in the quality and productivity of learning and teaching in universities?

By this we mean the capacity to determine in valid and reliable ways whether the teaching is better in one course, school, faculty, university than another (a workable distinction in practice—Drengenberg & Bain, 2016). We contend that at present the standards do not have such a capability. Most obviously, it is difficult to set discriminating standards for processes and practices that have not been articulated at scale. The lack of professionally controlled practice results in a retreat to generic and ambiguous terms in standards that lack the kind of discriminable quality required to generate workable distinctions. We will demonstrate this by way of example in our presentation and show how easily this circumstance can be altered.

Question 3:

Can we determine the quality of university learning and teaching at scale?

Without an understanding of what learning and teaching means at scale and the ability to set standards capable of producing workable distinctions in practice, it comes as no surprise that making a determination of quality is also exceedingly difficulty. According to NRC, (2012) when reporting the result of the work of the (US) National Council on Improving Measurement of Productivity in Higher Education:

“The panel looked carefully at the prospects for developing the kinds of comprehensive learning quality measures needed...We would have liked nothing better than to propose such measures but, unfortunately, we were forced to conclude that this will not be possible anytime soon.”

This determination included the consideration of the much-used student evaluation of teaching measures employed by most universities to determine the quality of their teaching. The fact that such measures do not stack up as a legitimate inclusion in productivity measures comes as no surprise given the previous remarks about the field’s inability to establish professional control over its practice. Without such control, students when completing evaluations are simply expressing satisfaction with a loosely defined articulation of a learning experience that is largely constructed idiosyncratically. It is not surprising that a nationally endorsed panel would find such an approach to be an insufficiently robust representation of the quality of teaching. According to Massy (2016), “Educational quality is the ‘elephant in the room’ in most discussions of higher education productivity”.

Carrying on Regardless: “*Measuring What We Have*”

Alarming, the inability to establish professionally controlled practice at scale and determine quality in higher education learning and teaching has not inhibited efforts to make quality and productivity judgments through productivity indices, institutional rankings, learning analytics, and an array of connoisseurship-type approaches.

Productivity Indices

In economic terms, productivity is a measure of the efficiency of a person, machine, factory, system, etc., in converting inputs into useful outputs. Productivity is computed by dividing average output per period by the total costs incurred or resources (capital, energy, material, personnel) consumed in that period (businessdictionary, 2016). The computation of productivity has been applied to higher education in the form of single and multi-factor productivity ratios (Massy, 2016). Single factor productivity models involve a single input and output measure; for example, relating student credit hours to staff full-time teaching load equivalency (FTE). Multi-factor models broaden the scope of inputs and outputs. There is a sense in the literature that multifactor models become more robust the more they can drill down on the nature of teaching activity (Massy, 2016). For example, how much time faculty spend developing lectures or online learning experiences and using resources.

From our perspective, the time taken to develop a lecture or learning experience is subordinate to the effectiveness of those activities in design and delivery using professionally controlled practice to maximize student outcomes. The true value of preparation time is not about whether it happens or the time it takes but what transpires as a result. Imagine determining the productivity of a medical or legal practice by establishing whether patients or clients are seen, hours/appointments billed, and resources consumed, without knowing whether the use of time and consumption of resources results in the successful resolution of cases or whether patients are regaining their health.

As bizarre as this approach may seem, the analogy is an accurate representation of what is occurring in university productivity measurement right now. The widespread and growing pressure related to the escalating costs of higher education and questions about the broader social return on the investment is applying immense pressure to universities to demonstrate they have a handle on their productivity. In response to this productivity pressure, universities are in a “full steam ahead- damn the torpedoes” frame of mind as they acquire software applications that model inputs and outputs and predict future trends sharpening the understanding of their operating costs, and making predictions of future performance. In this full-on rush to manage risk and be more accountable, the quality and efficacy of their core activity is simply assumed.

Remarkably, under such circumstances it is altogether possible for a university to improve its handle on learning and teaching “productivity” without knowing what learning and teaching or learning and teaching quality means at any scale-or even recognizing the need to ask that question? We find it hard to imagine how a university could reasonably control cost, for example without a more fulsome understanding of what it is doing and whether it is of quality.

Rankings

Despite the enormity of the problem we are describing, the sector persists in measuring what it has even if, as we are indicating, it is not what it needs. This is evident in the persistence with the “3 R” approaches to asserting quality: **R**ankings, **R**esearch and **R**eal Estate (Craig, 2015) which have little to do with the quality or productivity of learning and teaching, yet have become the desperate fodder of strategic plans and KPIs as leaders seek to move their institutions up the ladder of perceived national and international excellence. The marketing of university’s learning and teaching to prospective students based on rankings is at best misleading if not disingenuous especially when a university has so little understanding of and control over the way learning and teaching occurs within the institution.

Learning Analytics

Learning analytics have been widely adopted as a solution to the learning and teaching measurement problem. “Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs” (van Harmelen & Workman, 2012, p.7).

The problem here is that learning analytics data do not measure learning or teaching. When other fields use analytics systems to watch the behavior of shoppers, baseballers, fraudsters, or stock traders, they are watching what people do when they shop, play baseball, commit frauds or make trades. These are the behaviors the analysts are interested in understanding, predicting, and ultimately influencing. In learning analytics, the basic intent is the same. However, knowing about how many times a student uploaded a document, tweeted, or used a chat room has little to do with the ways students learn or how teachers teach. Further, making this data bigger by gathering lots of it does not make it any better in terms of measuring learning and teaching. Even more problematic is the use of this data to label students or make decisions about the promotion or termination of faculty members. Learning analytics like university rankings are surrogates reflecting the inability to answer the questions about professional control and develop standards capable of discerning quality.

Connoisseurship

A fourth example of the “*measuring what we have*” approach is the widely employed process of connoisseurship through peer review in higher education. In its contemporary form, peer panels are invited to cast judgment about the quality of an assessment regime or an institution’s learning and teaching governance approach. It is not so much that peer review is fundamentally problematic, but the issues and questions that transpire when such a process is not guided by standards capable of producing workable distinctions in learning and teaching practice. Absent real standards, such an approach is more like the kind of activity traditionally associated with the term (e.g., wine tasting or art critique) and less like the empirical evaluative approach most likely desired for determining the operational quality and

productivity of multi-million dollar public institutions. We view this type of connoisseurship as yet another example of surrogacy for valid and reliable measurement of professionally controlled practice. Peer review can be a powerful model for evaluation and accreditation when those doing the reviewing have the expertise and are equipped with the standards required to drive empirically defensible workable distinctions in quality— when the review is not merely subjective but carried out within an agreed framework, itself based upon empirical demonstration of actual outcomes.

Sameness and Disintermediation

There can be no doubting the potential benefit a university could accrue if it was capable of showing the effect of its learning and teaching approach on student learning outcomes at scale. Understanding the relationship between learning and teaching inputs and outputs (its quality and productivity) could potentially enable university learning and teaching to become ***better, deeper, faster, and more affordable***. A university capable of differentiated performance in any of the four areas would definitely have a distinctive story to tell.

The fact this is not currently the case may be best evidenced by the profound sameness in the way universities message their learning and teaching excellence. Mission documents and annual reports are replete with descriptions of teaching award winners, grant recipients, technological capability, and teaching facilities liberally sprinkled with claims about being “*world class*,” “*state of the art*”, and “*sector leading*.” That this is the best we can do in representing learning and teaching is an unfortunate consequence of the “*measuring what you have over what you need*” problem and an excellent example of the face validity of the case we are making in this presentation. All institutions claim a uniqueness of mission and value proposition for students yet their reality may be best represented by Dvorak & Busted (2015) who suggest that “the lack of enduring and unique identities in higher education offers up an opportunity for education leaders, as it indicates there are a host of undifferentiated brands ripe for disintermediation.” We should also note here that we reject many of the binaries in education discourse, in this case around education *versus* running a business (and also the binary of running the business to make money, so as to be able to do education). As in other fields, there is an unquestionable need to be able to show how the everyday, core professional transactions are *at the same time* integral to the funding and accountability mechanisms of the work. This is a basic requirement of any evidence-based practice.

So, what to Do?

Emergent Standards

We believe standards and standards agencies have an important role to play in responding to the current problems associated with measuring quality and productivity in higher education. At present, we believe the standards frameworks governing higher education operate on the assumption that universities are in possession of the models of practice, workable distinctions, and understanding of quality and productivity at scale that simply do not exist. The result is at best a set of quality judgments based on correlates of success over actual performance. A much better approach would be to employ standards in an *emergent* way to drive resolutions to the problems described in this paper. This involves setting standards that drive the sector to create the models of practice, learning and teaching capability, technologies and governance processes that could make measurement of quality and productivity legitimate based on a demonstrable relationship between university teaching and student learning. In other words, such standards would be *navigational* in telling those who use them that if X occurs, then Y should follow. Rather than the current situation as we described earlier, where education standards currently lack above all else any sort of detailed discrimination of everyday work. At present, we are engaged in a largely unproductive effort impersonating fields that possess the visible and comparable professionally controlled practice capable of producing workable distinctions in quality. We will demonstrate in our presentation what emergent standards look like and how they can enable professionally controlled practice.

Different Technologies and Genuine Learning and Teaching Analytics

An *emergent standards* focus could catalyze the development of technologies that do more than automate learning and teaching practice (as is the case with contemporary learning management systems) to help faculty design and deliver learning experiences that are capable of producing efficacious outcomes for students. These technologies could generate the kind of genuine “learning and teaching” analytic data capable of driving quality improvement. Such data would reflect the measurement of things known to make a difference for students and reflect workable distinctions in quality and productivity. This is also the kind of emergent feedback data (Bain, 2007; Bain & Drengenberg, 2016) required to build the whole-of-organizational big learning and teaching data sets that show how a university adds value to the learning experience of students. We will show an example of this technology in our presentation.

Governance and Peer Review

Emergent Standards that require universities to build policy and governance processes focused upon quality learning and teaching practice, could transform governance models from their well-documented focus the administrative correlates of learning and teaching (Zundans-Fraser & Bain, 2015) to the quality and productivity of comparable and visible professionally controlled practice. *Emergent standards* could transform the way universities are structured to support their core activity helping faculty members to do the normal work of learning and teaching because what those faculty members are doing is known and understood at organizational scale.

Conclusion

We will close with two summative assertions based upon the case made in this presentation. They are:

Productivity in Higher Education cannot be measured without an understanding of Quality.

Quality cannot be determined without an understanding of Professionally Controlled Practice at Scale.

Higher education is not the first branch of the education family tree to get caught up modeling its processes for determining quality and productivity on more mature fields that already possess a foundation of accepted professionally controlled practice at scale. Education as a whole has a long history of adopting/impersonating systems, tools and methods that have transformed the core activity of other fields, yet languish when applied in ours. The field's experience with information technology offers up a plethora of recent well documented examples in support of this claim. We believe it is time to take up the quality and productivity question and its implications seriously. This means first creating the conditions required to measure quality and productivity legitimately. Our concluding question in this presentation is to our community: *Is it not time to look at how we can employ standards to drive legitimate change in our organizations' approaches to learning and teaching at scale?* And as a consequence, align the measurement of what we have with what we need to demonstrate in defensible ways the quality and productivity of what we are doing in university learning and teaching.

References

- ACT (2009). *ACT report: Voluntary system of accountability learning gains methodology*. Retrieved from https://cp-files.s3.amazonaws.com/22/ACTReport_LearningGainsMethodology.pdf
- Bain, A. (2007). *The self-organizing school. Next generation comprehensive school reforms*. Lanham, MD: Rowman & Littlefield.
- Bain, A., & Drengenberg, N. (2016). *Transforming the measurement of learning and teaching in higher education*. New York: Routledge
- Bowker, G. & Starr, S. (2000). *Classification and its consequences*. Cambridge, Massachusetts: The

MIT Press.

- Businessdictionary (2016). Productivity. Retrieved from <http://www.businessdictionary.com/definition/productivity.html>
- Craig, R. (2015). The Problem With College Rankings. Retrieved from <https://www.higheredjobs.com/blog/postDisplay.cfm?post=704>
- Drengenberg, N., & Bain, A. (2016): If all you have is a hammer, everything begins to look like a nail – how wicked is the problem of measuring productivity in higher education? *Higher Education Research & Development*, DOI: 10.1080/07294360.2016.1208640
- Dvorak, N., & Busted, B. (2015). Its hard to differentiate one higher-ed brand from another. *Gallup Business Journal*, Retrieved from <http://www.gallup.com/businessjournal/184538/hard-differentiate-one-higher-brand.aspx>
- Liu, O. (2008). *Measuring learning outcomes in higher education using the measure of academic proficiency and progress (MAPP)*. Retrieved from <https://cp-files.s3.amazonaws.com/23/RR-0847MeasuringLearningOutcomesUsingMAPP.pdf>
- Massy, W. (2016). *Reengineering the university. How to be mission centered, market smart and margin conscious*. Baltimore: Johns Hopkins University Press.
- Mawdsley, R., & Cumming, J. (2008). Educational malpractice and setting damages for ineffective teaching: A comparison of legal principles in the USA, England, and Australia. *Education and the Law (20)*1, 25–46.
- (NRC) National Research Council (2012). *Improving measurement of productivity in higher education. Panel on Measuring Higher Education Productivity: Conceptual Framework and Data Needs*. T. A. Sullivan, C. Mackie, W. F. Massy, & E. Sinha (Eds.). Committee on National Statistics and Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Pascarella, E., & Terenzini, P. (2005). *How college affects students (Volume 2): A third decade of research*. San Francisco: Jossey-Bass.
- Scott, G., Coates, H., & Anderson, M. (2008). *Learning leaders in times of change: Academic leadership capabilities for Australian higher education*. Retrieved from http://research.acer.edu.au/cgi/viewcontent.cgi?article=1001&context=higher_education
- Scott, D., & Scott, S., (2015). Leadership for quality university teaching: How bottom-up academic insights can inform top-down leadership. *Educational Management Administration & Leadership*, 1-21 DOI: 10.1177/1741143214549970
- Zundans-Fraser, L. & Bain, A. (2015). How do institutional practices for course design and review address areas of need in higher education? *Higher Education Research & Development*, 35(4), 841-853, DOI: 10.1080/07294360.2015.1137883

