



Heriot-Watt University
Research Gateway

Quantifying Learning

Citation for published version:

Kandiko Howson, C & Buckley, A 2020, 'Quantifying Learning: Measuring Student Outcomes in Higher Education in England', *Politics and Governance*, vol. 8, no. 2, pp. 6-14.
<https://doi.org/10.17645/pag.v8i2.2564>

Digital Object Identifier (DOI):

[10.17645/pag.v8i2.2564](https://doi.org/10.17645/pag.v8i2.2564)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Politics and Governance

Publisher Rights Statement:

© Camille Kandiko Howson, Alex Buckley.

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Article

Quantifying Learning: Measuring Student Outcomes in Higher Education in England

Camille Kandiko Howson ^{1,*} and Alex Buckley ²

¹ Centre for Higher Education Research and Scholarship, Imperial College London, London, SW7 2AZ, UK;
E-Mail: c.howson@imperial.ac.uk

² Learning and Teaching Academy, Heriot-Watt University, Edinburgh, EH14 4AS, UK; E-Mail: alex.buckley@hw.ac.uk

* Corresponding author

Submitted: 17 October 2019 | Accepted: 25 November 2019 | Published: 9 April 2020

Abstract

Since 2014, the government in England has undertaken a programme of work to explore the measurement of learning gain in undergraduate education. This is part of a wider neoliberal agenda to create a market in higher education, with student outcomes featuring as a key construct of value for money. The Higher Education Funding Council for England (subsequently dismantled) invested £4 million in funding 13 pilot projects to develop and test instruments and methods for measuring learning gain, with approaches largely borrowed from the US. Whilst measures with validity in specific disciplinary or institutional contexts were developed, a robust single instrument or measure has failed to emerge. The attempt to quantify learning represented by this initiative should spark debate about the rationale for quantification—whether it is for accountability, measuring performance, assuring quality or for the enhancement of teaching, learning and the student experience. It also raises profound questions about who defines the purpose of higher education; and whether it is those inside or outside of the academy who have the authority to decide the key learning outcomes of higher education. This article argues that in focusing on the largely technical aspects of the quantification of learning, government-funded attempts in England to measure learning gain have overlooked fundamental questions about the aims and values of higher education. Moreover, this search for a measure of learning gain represents the attempt to use quantification to legitimize the authority to define quality and appropriate outcomes in higher education.

Keywords

accountability; education; governance; learning; quality assurance

Issue

This article is part of the issue “Quantifying Higher Education: Governing Universities and Academics by Numbers” edited by Maarten Hillebrandt (University of Helsinki, Finland) and Michael Huber (University of Bielefeld, Germany).

© 2020 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Since 2014, the government in England has undertaken a programme of work to explore the measurement of learning gain in undergraduate higher education, defined for the purposes of the programme as “a change in knowledge, skills, work-readiness and personal development, as well as enhancement of specific practices and outcomes in defined disciplinary and institutional contexts” (Kandiko Howson, 2019, p. 5). This is part of a wider neoliberal agenda in England, as over the past

decade the government has driven the development of a competitive market in higher education (Naidoo & Williams, 2015; Olssen, 2016). Browne (2010) suggested new forms of financing higher education and supporting widening participation (2010), with the Department for Business, Innovation and Skills moving to put ‘students at the heart of the system’ through shifting the burden of funding more completely from grants to tuition fees, and from the state to students (2011); home student fees trebled to £9,000 per year in 2012 under the leadership of the Minister of State for Universities and

Science David Willetts. A competitive market was fully put in place through the removal of student number allocation and the complete uncapping of student numbers by the Treasury in 2015. A market for students—with associated neoliberal ideology of a subsequent increase in quality—was designed, linking teaching excellence, social mobility and student choice (Department for Business, Innovation and Skills, 2016). This was implemented through new managerialism within higher education with a focus on outputs such as rankings to drive competition within a neoliberal market (Lynch, 2015).

Under neoliberal logic, to support a competitive market there is a need for information on how institutions are performing. Given the thousands of courses across hundreds of diverse institutions there is intense subjectivity in how ‘excellence’ is understood; however, quantification of performance gives the “appearance of scientific objectivity” (Ehrenberg, 2003, p. 147). This provides rankings and frameworks with their credibility as resources of information and as arbiters of value for higher education.

This neoliberal agenda understands ‘value’ primarily in terms of “corporate culture” and individual monetary gain (Giroux, 2002, p. 429), with student outcomes featuring as a key construct of value for money for students, alongside value for money for the state. These notions of value are increasingly subjected to measurement. However, a perennial question of social science research remains: are those meaningful concepts of value? And if they are not, what is the value of the measure? The assessment of learning gain started as a debate about the benefits that students were accruing from their time and investment in higher education. However, those more fundamental questions about quality have been lost in a search for quantity—the need for a numerical representation of quality, even if divorced from what it represents. In this article we explore the issues raised by the process of quantification represented by the learning gain initiative, particularly around who decides what students should learn, what higher education is for and how its value is measured. We suggest that the recent search for measures of learning gain in the UK is an example of a shift from quantification as a mechanism for representing value, to quantification becoming the value itself.

2. Interest in Large-Scale Learning Metrics

A range of evidence has prompted concerns about the value of what students derive from their investment in higher education, mostly out of the US due to escalating tuition fees and practices of for-profit providers. Research from the US indicates that there is a gap between employers and graduates’ views on the level of achievement of essential employability skills (Hart Research Associates, 2015), and varying conceptions of employability skills across stakeholders (Tymon, 2013). There is debate about the role of using employability metrics in higher education outcomes, particularly in re-

lation to generic outcomes as employers often have specific skill requirements from graduates (Cranmer, 2006; Frankham, 2016). A high-profile study in the US using the Collegiate Learning Assessment (CLA) instrument to explore what students are gaining from higher education seemed to find that significant proportions of students are not developing key skills such as critical thinking and complex reasoning (Arum & Roska, 2011). This raised questions about what students were learning and whether it was ‘enough.’

This question was at the heart of an Organisation for Economic Co-operation and Development feasibility study, the Assessment of Learning Outcomes in Higher Education. It was run across multiple countries and subjects of study. However, it faced challenges around questions of what to measure, with international, cultural and subject-level differences emerging. Due to concerns about data quality and use, the project was not continued (Organisation for Economic Co-operation and Development, 2013a, 2013b). This project identified the challenge of trying to develop a generic instrument across different disciplinary and national contexts.

The findings from the US and questions being asked globally resonated in the UK, which faced extensive political debates and student protests about raising tuition fees, alongside concerns about ‘grade inflation’ promoted by rises in the awarding of first-class degrees (Bachan, 2017). As a complement to changing the funding system to promote a market culture in higher education, the Minister David Willetts identified a need for comparable information to promote student choice and for accountability of the large sums of student fees entering the system, backed by public loans.

Existing global rankings such as those produced by the Times Higher Education use quantification as the basis of quality, (Hazelkorn, 2015) but focus on research and reputation. In the UK, the domestic rankings, compiled by major newspapers, include measures of student satisfaction drawn from the National Student Survey. However, the National Student Survey does not attempt to directly measure student learning, and there has been very little effort to establish a correlation between the National Student Survey scores and successful learning; a rare recent study suggests they may in fact be inversely related (Rienties & Toetel, 2016).

In the 1990s in the US, a similar lack of large-scale data related to student learning was noted, alongside the rising importance of research and reputation-based rankings. This led to development of the National Survey of Student Engagement, which is a distillation of decades of evidence on what activities promote student success (retention, progression and completion) into items which provide actionable data for students and staff (e.g., asking questions in class, such as ‘Do students do this?’ or ‘Can staff provide more opportunity for this to happen?’). It also provides benchmarked data and has a well-developed evidence-base for enhancing teaching and promoting student learning. It is now used across

the world (Coates & McCormick, 2014), and although a version has been developed for use in the UK (Kandiko Howson & Buckley, 2017), it has had relatively limited impact due to competition from the nationally-mandated National Student Survey.

The challenges encountered by international efforts to measure student learning, and associated outcomes such as graduate employability, show the dominance of national issues in higher education policy making. Even when schemes such as the UK's Research Excellence Framework are adopted by other countries the policies are adapted locally and not used comparatively (see the Excellence in Research for Australia, 2018). Efforts to measure student learning are bounded by cultural, structural and institutional differences across countries. Different conceptual definitions and student populations mean many data elements are not comparative (Matsudaira, 2016). For example, international students are variously seen in a deficit model, as taking local places, as a drain on public services or as a financial benefit (see Kandiko Howson & Weyers, 2013). Without international benchmarks in place, however, national efforts to measure student learning are highly politicised, as they are costly to design and administer. To justify the substantial investment, initiatives need to show the value both of the development of measurement tools, and—for political reasons—of national higher education sectors.

3. Origin of Measures of Learning Gain in England

Through the political desire to create a competitive market in higher education (Department for Business, Innovation and Skills, 2011), the actions of various policy actors and global influences, the government in England embarked on a large-scale effort to measure student learning gain. The initial catalyst for the learning gain agenda was the changes to tuition fee structure and the identification by the Minister of a lack of information for students to make 'value' decisions about what and where to study. As an indication of the policy complexity, the work was originally driven by three sector bodies that no longer exist: the Department for Business, Innovation and Skills (whose University remit moved to the Department for Education in 2016) alongside the Higher Education Funding Council for England (whose activities were taken over by the new regulator, the Office for Students in 2018) and the Higher Education Academy (which merged into AdvanceHE in 2018). Work started with a scoping study which developed a definition of learning gain as "the 'distance travelled' or the difference between the skills, competencies, content knowledge and personal development demonstrated by students at two points in time" (McGrath, Guerin, Harte, Frearson, & Manville, 2015, p. xi). This broad, generic view of learning gain contrasted with the academic literature, which defines it more narrowly, for instance as "the academic and personal transferable attributes gained as a result

of the active pursuit of content-specific knowledge in a given course of study" (Coates & Mahat, 2014, p. 17).

In 2015, the Higher Education Funding Council for England then led on designing three strands of activity to test various methodological approaches to measuring learning gain. Firstly, there was a suite of 13 pilot projects involving over 70 institutions. A second area focused on analysis of existing government databases to explore the possibility of finding proxy measures of learning gain. The third strand was initially mooted as developing a standardised assessment for students, however after backlash from the sector this was reconceptualised as a project based on the Wabash National Study led by the Center of Inquiry (2016). The Wabash project was a large longitudinal study which used multiple process and output measures to explore the impact of liberal arts education on student learning across multiple institutions in the US (Pascarella & Blaich, 2013).

However, as the strands developed, it was not clear to stakeholders what was being measured, or why, compounded by changes at the Ministerial level which resulted in a lack of intellectual leadership of the agenda. The Higher Education Funding Council for England provided an amended definition of learning gain on its website when the projects were launched, as "an attempt to measure the improvement in knowledge, skills, work-readiness and personal development made by students during their time spent in higher education" (2018, p. 1). Most of the pilot projects developed their own working definition of learning gain, referenced in project web-pages (Higher Education Funding Council for England, 2018), such as The Open University-led project adopting "a growth or change in knowledge, skills, and abilities over time that can be linked to the desired learning outcomes or learning goals of the course" the University of Lincoln-led project using "the extent to which undergraduate students have gained a key set of transferable skills and competencies that prepare them for the next stages of their career upon graduation, be it employment or further study" and "the extent to which participating in work-based learning, or work preparation activities, contributes to the readiness of the graduate to participate in a professional context" by the Ravensbourne-led project. These varied definitions indicate the complex territory of learning gain and the lack of consensus over what 'counts' as learning gain; measures are not neutral; they define what matters (Lynch, 2015; Power, 1994).

This led to debate across the sector about what constitutes a learning gain measure, with learning gain becoming an umbrella term for a wide variety of indicators relating to the student experience and student outcomes. There was further confusion with the development of the Teaching Excellence Framework, led by the Department for Business, Innovation and Skills, which aimed to assess teaching excellence and to adopt principles of quality-based funding, with 'Student Outcomes and Learning Gain' as one of the three pillars of quality explored (Gunn, 2018). Although technically separate policy initiatives,

there was extensive speculation in whether the learning gain programme would develop an outcomes metric that could be used for institutional comparison linked to funding. Furthermore, when taking over from the Higher Education Funding Council for England part-way through the learning gain projects, the Office for Students set itself up as a data-driven regulator, but without a clear position on future plans for learning gain.

Due to a lack of leadership of the initiative, the various sector stakeholders could not agree whether a use for the metrics should come first, such as designing institutionally comparative measures to measure performance and provide accountability, or whether valid measures of learning gain needed to be developed, that then could potentially be used for a variety of purposes, including enhancing teaching and learning and assuring quality. The projects struggled to develop measures without a clear direction for what they would be used for, as this impacts how measures are designed.

Whilst valid measures in specific disciplinary or institutional contexts were developed, such as concept inventories in Chemistry and mathematical models for institutions delivering higher education in further education settings (Kandiko Howson, 2019), a robust single instrument or measure failed to emerge. It also became apparent that the metrics devised were not as straightforward as hoped for by policymakers. Even existing measures such as students' grades demonstrated wide discrepancies across modules, courses and institutions.

The programme of work was beset with challenges of student engagement and interrelated issues around data protection, data sharing and research ethics. These challenges stemmed from a lack of rationale or clear purpose for measuring and using the data. Indeed, "The greatest challenge in developing learning indicators is getting consensus on what kind of learning should be measured and for what purpose a learning indicator is to be used" (Shavelson, Zlatkin-Troitschanskaia, & Mariño, 2018, p. 251). The focus on developing measures, rather than what needs measuring and why, has led to a circular policy development model rather the usual uni-directional causal model (Birkland, 2015). The outcomes of the learning gain programme became a solution in search of a problem. They successfully identified disciplinary-level differences both in terms of absolute outcomes but also in terms of what was valued, such as what successful communication skills are in Medicine and Law, and the role of reflection in Humanities and pre-professional subjects. However, government policy and regulatory levers operate instead at the institutional level.

4. Learning Gain and the Disciplines: US and UK Examples

The projects identified the discipline as the primary unit of comparison for student learning outcomes in England. However, policymakers were interested in a generic in-

strument which could be used to compare institutions, which became the focus of the two other strands of activity in the programme. This has been a recurring dream in the UK (Yorke, 2008) but efforts to do so have been largely centred on the US (McGrath et al., 2015).

Part of the reason for this is that the nature of US higher education makes it more realistic to search for broad agreement about which learning outcomes are most important. Firstly, the widespread focus on general education in undergraduate programmes generates consensus about learning outcomes. For example, Arum and Roksa (2011) justify their use of the CLA in their influential study on the plausible grounds that there is a common acceptance among US institutions about the importance of general critical thinking and related general skills, reflected in periodic calls for comparative student outcome measures to be used in the accreditation process (Ewell, 2015). Secondly, there are well-defined groups of institutions who broadly agree about key learning outcomes. The liberal arts colleges are the best example of this, having an explicit focus on a broad-based education and the development of general attributes such as written and oral communication, critical thinking and ethical reasoning (Association of American Colleges and Universities, 2005). These common goals of liberal arts institutions allowed the Wabash study to meaningfully administer a range of instruments assessing students' general skills, including critical thinking and moral reasoning (Pascarella & Blaich, 2013).

However, unlike the US, English higher education does not have an explicit focus on general education. Students may take a small number of broader 'elective' classes, but nearly all of their time will be spent studying within a relatively narrow field (or two narrow fields, in the case of joint programmes). For example, students at Harvard are currently only required to take 56 of 128 credits in their subject specialism over the four years of their degree (Harvard University, 2019). Most English students studying for single honours can have all of their credits in their subject specialism over the three years of their degree. Similarly, in the UK students almost always enter university on a programme with a specified subject specialism, whereas in the US students specify their specialisation after only one or two years of study. There is also a relatively high degree of specialisation in the English school system, with students typically leaving with qualifications in only three subjects. In the US, by contrast, has a broad-based secondary school curriculum and college entry is normally based on a student's SAT score, which measures general mathematical, reading and writing skills.

The development and assurance of learning outcomes in the UK are in line with this level of relative specialisation, as they are undertaken by the discipline communities themselves. The primary way of ensuring that institutions are assessing students in the 'right' way (both in terms of content and standard) is the external examining system, which is a process of peer-review in-

ternal to the discipline, largely devoid of comparable learning gain metrics. Professional disciplines often need to satisfy requirements placed on them by their professional bodies; again, this process is internal to the discipline. Non-disciplinary processes for determining and assuring learning outcomes—at institutional- or sector-level—are standardly at a very high level and are generally limited to checks that the appropriate discipline-level quality processes have been adhered to. Subject benchmarks, which are broad descriptions of what students should learn in a particular discipline, play a sector-level role and are owned by a sector-level body—the Quality Assurance Agency—but they are developed by representatives of the disciplinary communities. In England therefore, it is true to say that the system of checks and balances around the undergraduate curriculum assumes that the ultimate arbiters of what students should learn in their time in higher education are the disciplinary communities. Non-disciplinary agents (institutions, government and non-disciplinary sector bodies) have limited influence over learning outcomes, which is generally limited to ensuring that the relevant within-discipline processes have been followed.

Given the emphasis on discipline specialisation in England, efforts to mimic US developments of generic learning gain instruments are ambitious at best, and potentially misguided. In addition to differing structures of degrees in the two countries, the US efforts to measure learning gain were addressing different issues than the UK. Subsequently, ‘what’ was being, ‘why’ it was being measured and ‘how’ it was measured do not allow for straightforward policy transfer. However, political interest in a generic instrument led the UK to attempt to use the same methods as in the US, without thinking about the rationale underpinning the design and use of the metrics.

5. Disciplinary Learning in National Contexts

Despite a policy impetus, there are therefore a number of formidable obstacles to the development and use of generic instruments to measure learning gain in England. For example, general skills would need to be assessed in a generic instrument when students have learnt those skills almost entirely in disciplinary contexts. Even in the US with its traditional focus on general education, there is evidence that students’ performance on a generic instrument such as the CLA is influenced by their field of study (Arum & Roska, 2008). The explosive impact of the 2011 study by Arum and Roska was based partly on the finding that students from fields that do not emphasise reading and writing perform less well on the CLA. This is unsurprising: with the best will in the world, the challenge of devising a test of general skills that does not discriminate between a history student and a physics student is daunting.

However, the deeper challenge concerns the authority to decide what the key learning outcomes of higher

education are. The high-stakes measurement of learning gain requires fundamental decisions about what students are expected to learn. Very little in the structures of English higher education indicate that that is appropriate for non-disciplinary agents—government, regulator, funding body, quality agency—to make those determinations. As described above, English higher education treats disciplinary academic communities as the ultimate arbiters of what students should learn. This does not rule out the development of generic instruments to measure learning gain. A disciplinary community may decide that general skills (e.g., numerical reasoning) are among their important learning outcomes, and that those skills can be validly assessed using generic assessment tools. However, the structures of English higher education indicate that the decision would rest with the disciplinary community; no non-disciplinary agent could persuasively claim the authority to decide what students ought to learn.

The recent developments in the measurement of learning suggest the role of the disciplines in determining and assuring what students should be learning is under question. The attempt by sector-wide, non-disciplinary agents to create instruments to measure learning gain, and by doing so to implicitly claim authority over the key learning outcomes of higher education, fits with broader patterns of administrative and managerial encroachment on academic authority: 1) the more assertive behaviour of administrative agents (Bleiklie, 1998); 2) the more hands-on role of management (Deem, 2017), the usurpation of professional expertise by management expertise (Amaral, Meek, Larsen, & Lars, 2003) inspired by the reduction in trust in professional expertise (Beck & Young, 2005); and 3) the demystification of academic work in order to facilitate its management using generic tools and techniques (Henkel, 1997). The literature on managerialism in higher education focuses on the increasingly muscular presence of administrative and managerial units within institutions, but a parallel process has been occurring at sector-level, with organisations such as the Quality Assurance Agency and the Office for Students taking on increasing power within themselves at the expense of disciplinary communities (Becher & Trowler, 2001; Filippakou & Tapper, 2019). The amplification of the market in the English higher education system—increased fees, removal of number caps, introduction of ‘kitemarks’ via the judgements of the Teaching Excellence Framework—has coincided with encroachments on the responsibilities of academics, such as frequent accusations by (successive) higher education Ministers that they are failing to maintain appropriate standards and allowing ‘grade inflation.’

6. Learning Gain and the Purpose of Higher Education

The attempt to quantify learning raises questions about the purpose and underpinning values of higher education and necessitates debate about the rationale for

quantification—whether it is for accountability, measuring performance, assuring quality or for the enhancement of teaching, learning and the student experience. Metrics have many uses, but there is inherent tension between metrics used for accountability and improvement (Kuh & Ewell, 2010). Through focusing on ‘how’ to measure learning gain, the learning gain programme of work did not address the question of what quality is in higher education, or the more profound question of what higher education is for; the answers have a significant impact on the use of any resulting data. There is a ‘paradoxical tension’ between how academic staff and external stakeholders view accountability by student learning outcomes (Borden & Peters, 2014). The assumption that it is in the gift of government and sector-level funding bodies and regulators to define measures of learning gain usurps the authority of disciplines as the arbiters of student learning. The absence of student voices also raises questions about their role in determining what their educational experience is for (Klemenčič, 2018).

In terms of assuring quality, there has been a broad shift from process and programme evaluation to outcome evaluation (Harvey & Williams, 2010). For example, there is increasing emphasis on salary data (drawing on the Longitudinal Education Outcomes dataset) as a metric of educational quality (Office for Students, 2019a). When it comes to learning gain, the tension around who ‘owns’ the measures has implications for evaluating performance. As found across the pilot projects, disciplinary differences in marking present challenges of using outcome data for cross-subject and institutional comparisons (Ylonen, Gillespie, & Green, 2018). Sector bodies such as funding councils and the new regulator work at institutional level. However, unless metrics have resonance at the disciplinary level, where students experience higher education, they will fail to meet the ultimate aims of assuring and improving the experience of students, in addition to lacking the legitimacy conferred by disciplinary authority. Desire for comparable metrics leads to a focus on standardized outcome tests over instruments designed to support student learning and enhance teaching (Douglass, Thomson, & Zhao, 2012).

7. Quantification as an End in Itself

The search for comparable information about student learning has led to a focus on the ‘quantity’ of learning a student receives from their investment in higher education. This simplistic quantification of learning ignores the merit of the content and the process of learning. Any measure of learning gain would always be a proxy of the activity itself; however, without a clear purpose for measuring and quantifying learning the proxy measures become divorced from the underlying activity. Furthermore, through using proxy measures in high-stakes quality frameworks, they become targets in themselves. This has been seen through the use of proportion of top grades awarded in league tables, and the

recent rapid escalation in grades across the UK sector (Palfreyman, 2019). Similarly, in the US the use of admission rate and yield metrics (the ratio of admitted students and those that matriculate) have dramatically impacted admissions practices in the US (Monks & Ehrenberg, 1999).

A lack of a rationale, beyond the initial ministerial catalyst, for measuring learning gain beset the learning gain programme. In the pilot projects, academics worried about ‘unintended’ use of metrics or ‘non-disclosed intentions’ around their use. Several projects concluded they would rather err on the side of not producing national measures rather than developing them and then hoping they were used for ‘good’ educational purposes. When learning gain is separated from debates about purpose, it allows available numbers to be used as proxy measures, resulting in many higher education metrics that are divorced from causal effects of institutions (Matsudaira, 2016).

There are wide ranging consequences of using proxy measures, particularly for vulnerable and disadvantaged groups (O’Neil, 2017), such as through geographical measures of deprivation that ignore individual circumstances and algorithms that normalise explained and unexplained attainment gaps by ethnicity (Office for Students, 2019b, 2019c). Social inequalities are perpetuated through quality judgements based on institutional reputation, a key sorting and selection criterion for many employers (Hazelkorn, 2015). In response many employers now design in-house recruitment mechanisms. These are often methodologically flawed and burdensome tests, which creates high inefficiencies for employers and graduates (Keep & James, 2010). Furthermore, numbers as proxies become ends in themselves:

The net result is that ranks become naturalised, normalised and validated, through familiarity and ubiquitous citation, particularly through recitation as ‘facts’ in the media. Rankings, thus, attain an unwarranted truth status that makes them self-fulfilling by virtue of their persistence and existence. (Lynch, 2015, p. 198)

The quantification of learning can distil a complex activity to a number, but without a rationale for developing, selecting and using measures the number loses any sense of purpose or meaning and becomes an end in itself. Learning gain becomes another metric to be used for marketing purposes (Polkinghorne, Roushan, & Taylor, 2017). Additionally, as a data-driven regulator, the Office for Students has also set key performance indicators for itself, with a measure of learning gain being one its 26 ‘Measures of Success’ (Office for Students, 2019d), meaning that the regulator needs to develop a measure *for its own use*.

Despite the challenges described in this article, the measurement of learning gain has immense potential for enhancing quality and performance in higher education (Kuh & Jankowski, 2018; Shavelson et al., 2018). For

example, developing ‘quantity’ measures of quality facilitates policy drives for competition, transparency and accountability, which are unlikely to dissipate. In the search for valid measures of teaching quality, learning gain—particularly when used as the basis for calculating the ‘value added’ by institutions and programmes—has benefits over proxy metrics such as student satisfaction and salary data. Quantification approaches could also in principle help align various disciplinary-based quality approaches, addressing concerns around equity of experience and differential outcomes (Kandiko Howson & Mawer, 2013). However, through focusing on ‘how’ to measure learning gain independent of ‘why’ to measure it, or ‘what’ to measure, the creation of a robust higher education quality system with comparable student outcomes and clear evidence of value for money has been set back by these recent developments. With a quality system aligned to disciplines, yet a regulatory system that holds institutions to account, simple, straightforward measures of the quality of what students are gaining in higher education have not emerged. As long as the disciplines act as the arbiters of quality in education, a debateable position itself, the development of meaningful institutional-level measures will be challenging.

8. Conclusion

The search for data about learning gain provides an illustrative example of the ‘evaluative state’ in English higher education. Sector agencies engage in efforts to develop quantitative instruments in areas where they have no explicit claim to authority, relying on a general sense of the right of administrative and managerial agents to monitor the outcomes of higher education institutions. Logics inherent elsewhere in the system—about the awesome technical challenges in measuring learning gain across disciplines and institutions, about the unintended impact of quality metrics, about the tension between accountability and improvement, about the lack of apparent purchase that quantitative indicators of teaching quality have on student recruitment, about the role of disciplines in determining and assuring learning outcomes—are overridden by the quantitative rationale. Developments that assume particular answers to fundamental questions about the value of higher education take place without any explicit consideration of those questions. The answers are provided by the systems and structures that have particular perspectives—managerialism, quantification—built in. Higher education is full of contentious developments that adopt the logic of quantification without explicit discussion and undermine or usurp traditional disciplinary-based methods of quality assurance, accountability and regulation. The search for sector-wide measures of learning gain in English higher education provides a limit to governance by numbers, and an example of the overextension of the logic of quantification and a failure to turn ‘what’ students learn into ‘how much’ was gained.

Acknowledgments

We would like to thank Maarten Hillebrandt and Michael Huber for organizing the workshop and thematic issue and for the helpful feedback from three anonymous reviewers.

Conflict of Interests

The authors declare no conflict of interests.

References

- Amaral, A., Meek, V. L., Larsen, I. M., & Lars, W. (Eds.). (2003). *The higher education managerial revolution?* (Vol. 3). Berlin and Heidelberg: Springer Science + Business Media.
- Arum, R., & Roska, J. (2008). *Learning to reason and communicate in college: Initial report of findings from the CLA longitudinal study*. New York, NY: Social Science Research Council.
- Arum, R., & Roska, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.
- Association of American Colleges and Universities. (2005). *Liberal education outcomes: A preliminary report on student achievement in college*. Washington, DC: Association of American Colleges and Universities.
- Bachan, R. (2017). Grade inflation in UK higher education. *Studies in Higher Education, 42*(8), 1580–1600.
- Becher, T., & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the cultures of disciplines*. Buckingham: Open University Press.
- Beck, J., & Young, M. F. (2005). The assault on the professions and the restructuring of academic and professional identities: A Bernsteinian analysis. *British Journal of Sociology of Education, 26*(2), 183–197.
- Birkland, T. A. (2015). *An introduction to the policy process: Theories, concepts, and models of public policy making*. Abingdon: Routledge.
- Bleiklie, I. (1998). Justifying the evaluative state: New public management ideals in higher education. *Journal of Public Affairs Education, 4*(2), 87–100.
- Borden, V. M., & Peters, S. (2014). Faculty engagement in learning outcomes assessment. In H. Coates (Ed.), *Higher education learning outcomes assessment: International perspectives* (pp. 201–212). Bern: Peter Lang GmbH.
- Browne, J. (2010). *Securing a sustainable future for higher education: An independent review of higher education funding and student finance* (Report BIS/10/1208). London: Department for Business, Innovation and Skills.
- Center of Inquiry. (2016). Wabash national study 2006–2012. *Centre of Inquiry*. Retrieved from <https://centerofinquiry.org/wabash-national-study-of-liberal-arts-education>

- Coates, H., & Mahat, M. (2014). Advancing student learning outcomes. In H. Coates (Ed.), *Higher education learning outcomes assessment: International perspectives* (pp. 15–31). Bern: Peter Lang GmbH.
- Coates, H., & McCormick, A. (Eds.). (2014). *Engaging university students: International insights from system-wide studies*. London: Springer.
- Cranmer, S. (2006). Enhancing graduate employability: Best intentions and mixed outcomes. *Studies in Higher Education, 31*(2), 169–184.
- Deem, R. (2017). New managerialism in higher education. In J. C. Shin & P. Teixeira (Eds.), *Encyclopaedia of international higher education systems and institutions* (pp. 1–7). Dordrecht: Springer.
- Department for Business, Innovation and Skills. (2011). *Higher education: Students at the heart of the system*. London: Department for Business, Innovation and Skills.
- Department for Business, Innovation and Skills. (2016). *Success as a knowledge economy: Teaching excellence, social mobility & student choice*. London: Department for Business, Innovation and Skills.
- Douglass, J. A., Thomson, G., & Zhao, C. M. (2012). The learning outcomes race: The value of self-reported gains in large research universities. *Higher Education, 64*(3), 317–335.
- Ehrenberg, R. G. (2003). Reaching for the brass ring: The US News & World Report rankings and competition. *The Review of Higher Education, 26*(2), 145–162.
- Ewell, P. (2015). *Transforming institutional accreditation in US higher education*. Boulder, CO: National Center for Higher Education Management.
- Excellence in Research for Australia. (2018). Australian government. *Australian Research Council*. Retrieved from <https://www.arc.gov.au/excellence-research-australia/era-2018>
- Filippakou, O., & Tapper, T. (2019). The state, the market and the changing governance of higher education in England: From the University Grants Committee to the Office for Students. In O. Filippakou & T. Tapper (Eds.), *Creating the future? The 1960s new English universities* (pp. 111–121). Cham: Springer.
- Frankham, J. (2016). Employability and higher education: The follies of the ‘productivity challenge’ in the Teaching Excellence Framework. *Journal of Education Policy, 32*(5), 628–641.
- Giroux, H. (2002). Neoliberalism, corporate culture, and the promise of higher education: The university as a democratic public sphere. *Harvard Educational Review, 72*(4), 425–464.
- Gunn, A. (2018). The UK Teaching Excellence Framework (TEF): The development of a new transparency tool. In A. Curaj L. & R. Pricopie (Eds.), *European higher education area: The impact of past and future policies* (pp. 505–526). Cham: Springer.
- Hart Research Associates. (2015). *Falling short? College learning and career success*. Washington, DC: Association of American Colleges and Universities.
- Harvard University. (2019). Harvard University handbook for students 2019–2020. *Harvard University*. Retrieved from <https://handbook.fas.harvard.edu/book/welcome>
- Harvey, L., & Williams, J. (2010). Fifteen years of quality in higher education. *Quality in Higher Education, 16*(1), 3–36.
- Hazelkorn, E. (2015). *Rankings and the reshaping of higher education: The battle for world-class excellence*. Cham: Springer.
- Henkel, M. (1997). Academic values and the university as corporate enterprise. *Higher Education Quarterly, 51*(2), 134–143.
- Higher Education Funding Council for England. (2018). Learning gain. *Higher Education Funding Council for England*. Retrieved from <https://webarchive.nationalarchives.gov.uk/20180319113650/http://www.hefce.ac.uk/lt/lg>
- Kandiko Howson, C. B. (2019). *Final evaluation of the Office for Students learning gain pilot projects*. Bristol: Office for Students.
- Kandiko Howson, C. B., & Buckley, A. (2017). Development of the UK engagement survey. *Assessment & Evaluation in Higher Education, 42*(7), 1132–1144.
- Kandiko Howson, C. B., & Mawer, M. (2013). *Student expectations and perceptions of higher education*. London: King’s College London.
- Kandiko Howson, C. B., & Weyers, M. (Eds.). (2013). *The global student experience: An international and comparative analysis*. London: Routledge.
- Keep, E., & James, S. (2010). *Recruitment and selection: The great neglected topic* (SKOPE Research Paper 88). Cardiff: Cardiff University and SKOPE.
- Klemenčič, M. (2018). The student voice in quality assessment and improvement. In E. Hazelkorn, H. Coates, & A. C. McCormick, (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 332–346). Cheltenham: Edward Elgar Publishing.
- Kuh, G. D., & Ewell, P. T. (2010). The state of learning outcomes assessment in the United States. *Higher Education Management and Policy, 22*(1), 1–20.
- Kuh, G. D., & Jankowski, N. A. (2018). Assuring high-quality learning for all students: Lessons from the field. In E. Hazelkorn, H. Coates, & A. C. McCormick (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 305–320). Cheltenham: Edward Elgar Publishing.
- Lynch, K. (2015). Control by numbers: New managerialism and ranking in higher education. *Critical Studies in Education, 56*(2), 190–207.
- Matsudaira, J. (2016). Defining and measuring institutional quality in higher education. In K. Matchett, M. Lund Dahlberg, & T. Rudin (Eds.), *Quality in the undergraduate experience: What is it? How is it measured? Who decides?* (pp. 57–80). Washington, DC: National Academies Press.
- McGrath, C. H., Guerin, B., Harte, E., Frearson, M., &

- Manville, C. (2015). *Learning gain in HE*. Cambridge: Rand Corporation.
- Monks, J., & Ehrenberg, R. G. (1999). US News & World Report's college rankings: Why they do matter. *Change: The Magazine of Higher Learning*, 31(6), 42–51.
- Naidoo, R., & Williams, J. (2015). The neoliberal regime in English higher education: Charters, consumers and the erosion of the public good. *Critical Studies in Education*, 56(2), 208–223.
- O'Neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York, NY: Crown.
- Office for Students. (2019a). Graduate earnings data on Unistats from the Longitudinal Education Outcomes (LEO) data. *Office for Students*. Retrieved from <https://www.officeforstudents.org.uk/data-and-analysis/graduate-earnings-data-on-unistats>
- Office for Students. (2019b). Ethnicity. *Office for Students*. Retrieved from <https://www.officeforstudents.org.uk/advice-and-guidance/promoting-equal-opportunities/evaluation-and-effective-practice/ethnicity>
- Office for Students. (2019c). Young participation by area. *Office for Students* Retrieved from <https://www.officeforstudents.org.uk/data-and-analysis/young-participation-by-area>
- Office for Students. (2019d). Measures of our Success. *Office for Students* Retrieved from <https://www.officeforstudents.org.uk/about/measures-of-our-success>
- Olsen, M. (2016). Neoliberal competition in higher education today: Research, accountability and impact. *British Journal of Sociology of Education*, 37(1), 129–148.
- Organisation for Economic Co-operation and Development. (2013a). *Assessment of higher education learning outcomes* (Feasibility Study Report, Volume 2). Paris: Organisation for Economic Co-operation and Development.
- Organisation for Economic Co-operation and Development. (2013b). *Assessment of higher education learning outcomes* (Feasibility Study Report, Volume 3). Paris: Organisation for Economic Co-operation and Development.
- Palfreyman, D. (2019). Regulating higher education markets. In T. Strike, J. Nicholls, & J. Ruthforth (Eds.), *Governing higher education today: International perspectives* (pp. 202–216). London: Routledge.
- Pascarella, E., & Blaich, C. (2013). Lessons from the Wabash national study of liberal arts education. *Change*, 45(2), 6–15.
- Polkinghorne, M., Roushan, G., & Taylor, J. (2017). Considering the marketing of higher education: The role of student learning gain as a potential indicator of teaching quality. *Journal of Marketing for Higher Education*, 27(2), 213–232.
- Power, M. (1994). *The audit explosion*. London: Demos.
- Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior*, 60, 333–341.
- Shavelson, R. J., Zlatkin-Troitschanskaia, O., & Mariño, J. P. (2018). Performance indicators of learning in higher education institutions: An overview of the field. In E. Hazelkorn, H. Coates, & A. C. McCormick (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 249–263). Cheltenham: Edward Elgar.
- Tymon, A. (2013). The student perspective on employability. *Studies in Higher Education*, 38(6), 841–856.
- Ylonen, A., Gillespie, H., & Green, A. (2018). Disciplinary differences and other variations in assessment cultures in higher education: Exploring variability and inconsistencies in one university in England. *Assessment & Evaluation in Higher Education*, 43(6), 1009–1017.
- Yorke, M. (2008). *Grading student achievement in higher education: Signals and shortcomings*. Abingdon: Routledge.

About the Authors



Camille Kandiko Howson is Associate Professor of Education at the Centre for Higher Education Research and Scholarship at Imperial College London. She is an international expert in higher education research with a focus on student engagement; student outcomes and learning gain; quality, performance and accountability; and gender and prestige in academic work. She is a Principle Fellow of the Higher Education Academy.



Alex Buckley is an Assistant Professor in the Learning and Teaching Academy at Heriot Watt University in Edinburgh, Scotland. His work is focused on supporting individuals and groups to enhance learning and teaching. He has previously held roles at Strathclyde University and the Higher Education Academy (now AdvanceHE). His research interests include assessment and feedback, student engagement and student surveys.