Chapter 3

Learning Gain: Can It Be Measured?

Peter Gossman and Stephen Powell

3.1 Introduction: Measuring 'Learning Gain'

With an increasing tendency to see higher education as a product with a price tag, there is understandably growing interest in the extent to which academic programmes of study promote students' employability and earning power. (The Quality Assurance Agency for Higher Education (QAA), 2013, para. 1)

This chapter addresses the 'basics' underpinning the notion of learning gain including its measurement and the motivations behind the recent interest in quantifying (and attributing) changes in students, as brought about by their learning experiences at university. However, it is firstly worth considering a rather larger question: if we wish to quantify some gain in learning then arguably we first need to define what we are really seeking to measure, rather than starting from a position of what might actually be measured, in the hope that something can be found. This is a criticism that could perhaps be laid against the Higher Education Funding Council for England (HEFCE) learning gain programme (which will be addressed later). Put simply, learning gain becomes a different thing according to the instrument of measurement applied to it. Further, if we measure learning gain in one particular way, this says something about the kind of learning, and any gains within it, that we most value.

There are many stakeholders within the UK's higher education (HE) system, and when considering what to measure, it is important to recognise that there are different views about what the purpose of higher education is and, by extension, what we should measure to assess the effectiveness of particular approaches. For example, physicists may put great store in assessing key concepts that are core to becoming

P. Gossman (B)
University of Worcester, Worcester, UK
e-mail: p.gossman@worc.ac.uk
S. Powell
Manchester Metropolitan University, Manchester, UK
e-mail: stephen.powell@mmu.ac.uk

A. Diver (ed.), Employability via Higher Education: Sustainability as Scholarship,
https://doi.org/10.1007/978-3-030-26342-3_3

a physicist (Sands, Parker, Hedgeland, Jordan, & Galloway, 2018: 610): this might be done using inventory tools, but this is very different from assessing progress against broader curricular aims. For many students and employers, a state of workreadiness may be considered a core outcome of any university experience, although this is also a contested viewpoint (Cameron, Wharton, & Scally, 2018, p. 41). From a national funding perspective, one of the key purposes of measuring learning gain is to allow comparisons of the effectiveness of learning processes between different institutions, and to better understand what works and why. This is possibly motivated by a political desire to promote greater market competition (Blackmore & Kandiko, 2012). Professional associations, employers, parents, and students all have views on the purpose of a university education and, by inference, what learning gain should be measured. What are the implications of this for selecting the means of measurement? In an era of increasing institutional competition, could the sector come to agreement or will competing values result in a mixed economy of measurement determined by local needs and values?

It can be argued that the need to measure gain in HE is in fact due to the problem of not knowing what HE produces in terms of graduate outcomes, although many institutions have codified their expectations of what it means to be a graduate. Of course, this is not easy to answer simply, and HEFCE, in running a range of projects, notes that any valued gain exists in relation to a specific context, defining it as: 'The improvement in knowledge, skills, work-readiness and personal development made by students during higher education' (HEFCE, 2015). This broad definition does not discriminate between learning gain and the 'value added' aspects of education however (again, an issue that will be unpacked later in this chapter.) In contrast to a contextual approach, learning gain can instead be framed as a measure of accountability. However, this is perhaps a circular argument, since if students are prepared to pay (albeit in many contexts through soft loans) and commit to a significant amount of effort, then they must in some way value the product on offer. Nonetheless, as

Shavelson and Huang (2003) stressed, calls for assessment-based accountability are unlikely to go away. Similarly, the Organisation for Economic Cooperation and Development (OECD) has noted that:

In a complex, ever changing and growing higher education context, where a variety of rankings are often being used as the yardstick of academic excellence, there is a clear need for a way to effectively measure the actual outcomes of teaching and learning. (OECD, 2013, p. 2)

Shavelson and Huang (2003) also note that the diversity of HE institutions in the US (in terms of inputs, processes, outputs and valued outcomes) is vast, and that there are obvious parallels with the UK. The following section explores learning gain using Biesta's (2015) functions of education to propose a contextual framework that takes account of the plurality of valued outcomes and outputs of HEIs. We suggest that consideration must be given to the question of what institutions should measure in terms of learning, before jumping straight to the measurement of what can be easily measured: the latter approach has great potential both for distortion of results and is quite open to the effects of unintended consequences.

3.2 A Contextual Framework for Learning Gain?

Gossman, Powell and Neame (2018) use Biesta's (2015) three functions of education to consider learning gain in relation to institutions' mission statements around learning. These functions are: qualification, socialisation and subjectification. Some measures of learning gain will fit neatly into the 'qualification' category. 'Qualification,' is about an ability to do something:

a 'doing' that can range from the very specific (such as in the case of training for a particular job or profession, or the training of a particular skill or technique) to the much more general (such as an introduction to modern culture, or the teaching of life skills, etcetera). (Biesta, 2015, p. 19)

Immediately within this 'doing,' a wide range of learning gain measurements could be undertaken. HEFCE argue that an appropriate measure of gain is therefore one that is suited to the context of the institution as perhaps outlined in a mission statement. If we take a simple example, we could test a student on their ability to describe the role of 'x' and 'y' in the function of 'z'. The test could then be applied

at the start of a course of study and then repeated at the end. The difference between the results is the quantified learning gain in that aspect of 'qualification,' i.e. the one that we value. Biesta (2015) notes further that 'qualification' has a range of purposes that are not just economic insofar as this can also contribute to 'citizenship or cultural literacy more generally' (2015, p. 11). This serves as a further illustration of the debate surrounding educational purpose. Biesta's second function of education is socialisation: 'Through its socializing function education inserts individuals into existing ways of doing and being' (2015, p. 19). This function is perhaps most apparent when an institution adopts a particular ethos, or the course a student enrols upon relates to a set of professional values and practices that a student is then acculturated into (e.g. nursing or accountancy).

It is worth emphasising that whatever is identified as the key aspect of education to measure for learning gain also has a political context (aside from being valued, in itself). Clarke (2017) further notes that the 'pressure to provide evidence of the value and effectiveness of tertiary education has come from two main sources—governments and employer groups' (2017, p. 1). This statement has a bearing on both the skills within a given 'qualification,' and on the need for demonstrable learning gain in the first place. Put broadly, the assessment of HE learning outcomes (if we assume that learning gain can be defined and is measurable) can provide a comparison between individuals, and, when aggregated and benchmarked, between the courses studied, and the institutions where studying occurs. It is then a short step to league tables of learning gain and an overly simplified notion of greater gain equating to a better education and possibly better value for money.

The pressure, as noted above, from employer groups relates to graduate employability and the development of skills and attributes by students during their course: communication, teamworking, and problem solving (Sin & Neave, 2014). The range of titles for these skills, includes such terms as "transferable skills, generic skills, generic attributes, generic competences, generic capabilities or key skills" (Jääskelä,

Nykänen, & Tynjälä, 2016, p. 132) and is indicative of the scope of the field. As Jääskelä et al. note 'these terms refer to competencies that education should provide regardless of the specific field and that can be used in a variety of tasks' (2016, p. 132).

This is a key point, which raises the question of whether there is a set of competences that HE should provide? Operationalising concepts to enable these to be measured is a rich field for learning gain. See for example, the Wabash longitudinal study in the United States which examined the graduate outcomes of a liberal arts education (2006–2012).1 The selected outcomes are presented as follows:

• Critical thinking • Positive attitude toward literacy

• Moral reasoning • Interest in contributing to the arts

• Socially responsible leadership • Interest in contributing to the sciences

• Interest in engaging intellectually challenging work

• Openness to engaging new ideas and diverse people

• Interest in political and social involvement • Orientation toward interacting with diverse people

• Well-being • Academic motivation

As Biesta observed, aside from 'qualification' and 'socialisation,' the act of 'subjectification' is also an educational function. This process should 'allow those educated [to] become more autonomous and independent in their thinking and acting' (2015, p. 20). This has immediate appeal and it is hard to argue against this function being a key HE aim and/or outcome, with graduates gaining enhanced autonomy and the ability to engage in moral reasoning.

3.3 The Mechanics of Learning Gain

Implicit in learning gain and any measurement of it (especially if it is to be applied to individual students) is that it must involve two measurements, one at the start of a time-period—perhaps a course of study—and one at the end of it. Learning gain as such, for a student, is a simple idea and can be outlined as a pre/post test score difference (T1 to T2 position P1 to P4, as per Fig. 3.1). This appears to be straightforward and can be referred to as the 'distance travelled' by a student. However, in practice, in an HE context, it is rather more complex, especially if linked to benchmarked comparisons, where for example, any given student's gain is then compared to gains made by 'similar' students. Boud (2018) describes what would be required for assessment approaches to measure learning gain at a programme (or

course) level, which essentially comes down to the application of a consistent set

1Ernest T. Pascarella & Charles Blaich (2013) Lessons from the Wabash National Study of Liberal Arts Education, Change: The Magazine of Higher Learning, 45:2, 6–15, DOI: https://doi.org/10.1080/00091383.2013.764257.

Fig. 3.1 Visualisation of individual learning gain

of learning outcomes and graduated set of assessment criteria across the levels of a given programme. Although this approach could work at a programme level, it does nothing to address the desire for comparisons between institutions. A clear illustration of where the above approach works is in the compulsory education sector in England. There, a national curriculum framework sets out the programme of study and attainment targets for children from early years (age 3) to the end of Key Stage 4 (age 16). Initially the focus is on reading, writing and maths, before broadening out in Key stage 3 and 4 to a wider range of subjects. Because children are studying the same curriculum and have standardised assessments at the end of Key stage 1 and 2 (culminating for most in GCSE examinations at year 11), it is possible to measure the gain of individuals and groups of students and compare them against each other.2 In HE, there is no structured national curriculum nor are there national assessment standards beyond the QAA subject benchmark statements and standards set by professional regulatory bodies.

In Fig. 3.1 the performance of a student is measured by a test/assessment at the start of a programme or module (T1). There are many factors to be considered here. For example, how long will the assessment be, when will it be administered, could/should all students take it at the same time (if they do not, what variations in results might occur due to some unknown effect modifier?) and how many assessments will a student take? Following T1, the student undertakes their programme of study and

2See further https://www.compare-school-performance.service.gov.uk (accessed 10.11.18).

is re-tested at T2 on, for example, their teamwork skills. Again, there are issued to be considered: should the test be the same one, administered twice? Should it be a different one with the validity of the test assured in some way? At very least the

test must reliably address the same outcomes using consistent assessment criteria calibrated for the level of study. In Fig. 3.1 three gain results are shown: P1, P2 and P3 alongside one learning loss position (perhaps an unlikely outcome), P4. When the results of P1–P4 are aggregated across a programme of study and averaged (Pmean), the mean course/programme learning gain can be represented.

Further questions arise: to whom can the learning again be attributed? In the case of an individual, we might ask if the gain is solely that of the student? This would seem a logical answer, but gains might also be attributed to the teacher(s) of the programme, or perhaps to the student's parents or guardians, or to their friends and the encouragement they give. The average gain across a group of students might seem to be better attributed to the teachers, programme and institution, but is this really the case?

Taking this reasoning further, programme learning gains could in turn be aggregated together as a measure of institutional learning gain, as shown in Fig. 3.2.

However, the challenge of comparing learning gain between institutions rests upon the adoption of a common set of standards across the sector: everyone needs to be measuring the same set of things for this to work in a sensible way. Of course, this is not the case, and this is in part the reason why university league tables depend

Fig. 3.2 Visualisation of institutional learning gain

upon proxies for learning gain, such as employability that measure (arguably) the work-readiness aspects of learning gain (Cameron et al., 2018).

The next section highlights possible debates around what might be used as measures of learning gain and some of the challenges to be faced. For example, can learning gain for an institution simply be the aggregated and averaged difference between the two scores (at T1 and T2) for all that institution's students? At what point should pre and post-testing take place? Might different times for these tests mean different results for a student, or a particular group of students, or for different types of institutions? Can these mean scores for institutions be compared between different parts of an institution and across the sector? If this is not the case then pre/post score differences make less sense and a more benchmarked approach (with all the complexity that it brings in terms of predicting student progress on entry to

an institution and tracking their progression at exit from the institution) becomes more meaningful. This in turn moves learning gain closer to a 'value-added' notion. Whatever choices are made, there are significant resource implications. It is possible that an institution's mean learning gain score could be good when compared against the whole sector mean, but worse than the comparable benchmark intuitions' mean. It is also possible to hypothesise that high UCAS tariff institutions may not be able to demonstrate a greater distance travelled as compared with institutions with a lower entry tariff.

Before discussing the possible measure of learning gain however, it is worth exploring the distinction between measures of learning gain and value added. Learning gain, as illustrated above, involves an ipsative comparison. In other words, for an individual student, the T2 result is referenced back to a T1 result. In the case of value-added (see for example the A-Level Information System [Alis3]), the learning of the student is compared with a statistically predicted result. In the case of A-Level examinations, the grade of the A-level is assigned a numeric score and a student's actual result is compared with the statistically predicted result, extrapolated from GCSE results, and considering attitudinal data from the Alis questionnaire, which can in turn help understand a student's attitude to their learning environment. A range of other factors (such as postcode, socio-economic background or ethnicity) can be used in a context value added approach (CVA), to consider a wider range of variables that might impact upon learning. When the difference between actual and predicted results is positive then there is value-added and vice versa. Like learning gain, these individual scores can be aggregated and averaged, again for courses and institutions. And yet, there is a note of caution to be sounded here: the student population at universities can be highly selective, depending largely upon the filters of previous academic results. This means that not only will learners have different starting points but also that there will be concentrations of different ability levels within different institutions. In effect, this is like a setting in selective schools and

3'Alis' is a value-added measurement, described by Durham University's Centre for Evaluation and Monitoring (CEM) as 'providing detailed value-added progress information for each student and subject at the end of the course.' (CEM, 2017, para. 1). Alis was originally a straightforward value-added system: currently, more sophisticated versions are available whereby a test adapted

according to student responses can be used to provide a learner benchmark.

it would be expected that those student (who have proven already to have the most ability) would outperform their peers, having both higher levels of learning gain and apparently higher value added. However (as the Sutton Trust found), such

> …so-called 'grammar school effect' disappeared once they took into account the fact that pupils who entered grammar school were progressing faster than their peers before they entered the grammar schools… Theirs is a self-selected sample predisposed to prepare for and take entry tests and perform well, so they do much better than those of a similar general ability, (Wheadon, 2013, p. 1).

Arguably, to make more secure judgements about value-added, a sophisticated CVA approach is required.

3.4 What May Be Measured (HEFCE Learning Gain Projects)

At the time of writing, HEFCE (now The Office for Students)4 notes five potential ways of measuring learning gain. These measures stem from a RAND European report (McGrath, Guerin, Harte, Frearson, & Manville, 2015) in which 14 potential methods for measuring learning gain were reviewed and allocated to one of five ways. They are: grades, surveys, standardised tests, mixed methods, and other qualitative methods. Some of these are inspired by the potential for technical developments within the realm of learning analytics and linked data. Grades are a self-explanatory measure, but the lack of granularity in outcomes-based assessment as used in many HE systems causes significant issues (Boud, 2018). The five grade bands (1st class, 2:1 and so on) are too broad to illustrate gain between entry and exit for a student on any given course. In addition to this, there are serious challenges in attempting comparison across disciplines and institutions. In part, it is this lack of precision with grades that prompts some of the debate in relation to the requirements of illustrating learning gain. Self-reporting surveys by students report on gain extent in relation to what is asked, again, via an ipsative approach. For example, the National Student Survey (NSS) asks students about communication skills (Q.20).

As McGrath et al. (2015) note, this is only one of three potential learning gain

questions from the NSS: The United Kingdom Engagement Survey, (UKES) contains 12 skills development questions (see Table 3.1, for an example that relates directly to Biesta's (2015) educational purpose of socialisation).

Currently, the NSS and UKES are only undertaken once during a student's course of study. This would need to be addressed to allow T1 and T2 comparisons of gain.

Other possible surveys include: student skills audits, surveys of workplace skills and surveys of career readiness. In all cases the problems of self-reporting are apparent, as is the potential for such surveys to be 'gamed,' perhaps via the artificial depression of pre-test scores or some degree of 'coaching' of students by academic staff towards

4See The Office for Students (OfS) website https://www.officeforstudents.org.uk/advice-andguidance/teaching/learning-gain/ (accessed 09.11.18).

Table 3.1 UKES learning gain sample question

How much has your experience at this institution contributed to your knowledge, skills and personal development in (Q9) Being an informed and active citizen:

(Response categories)

very much / quite a bit / some / very little

desired responses that are mutually beneficial. Standardised tests can be generic (e.g. the Collegiate Learning Assessment test) or discipline-specific, and seek to measure gain in identified, stated skills. Discipline tests, as McGrath et al. (2015) note, can be more readily compared and thus have greater validity than generic ones. They give the example of repeatedly testing medical students during their studies using an exam of 125 items. Clearly testing like this is demanding of resources, even just in terms of generating questions, but it can provide longitudinal data for each student as well as school comparisons, so long as there is consistent use of the test instruments.

These first three of HEFCE's five categories for measuring learning gain (i.e. grades, standardised tests, and self-reporting surveys) can all be administered to follow the simple 'distance travelled' model (T1 to T2 position P1 to P4, as per Fig. 3.1).

Existing data collection instruments within institutions could be modified to make them suitable to this approach but in all cases there are significant administration costs. The least costly option might be via grades using a system similar to Alis.

However, this still fails to take account of context value added, an approach strongly

recommended by the Fisher Family Trust for Schools (Thompson, 2018). Mixed methods (as the name suggests) is the use of combined tools to track student performance changes. This draws on a range of tools and indicators to track improvements in performance (for example through a combination of grades, student learning data and student surveys). In terms of truly representing learning gain, such approaches arguably have the most to offer as they recognise the complexity of the challenge being faced. Balancing out the relative contributions of the different measures, and the resources required to administer such an approach, remain significant, however. It may be that in the medium and longer-term developments in the sophistication of learning analytics and the ability to more readily stitch together different data sets will make such approaches increasingly possible.

The fifth category, other qualitative methods, is rather more grounded in mastery of learning and is markedly different from pre/post difference. These encourage students to reflect upon their own learning, acquired skills, and any skills gaps. They can also stimulate productive discussions between students and their tutors. Arguably, these methods offer much value insofar as they not only assess learning but serve to promote it through reflective and evaluative processes.

In terms of research done in the U.S., Benjamin and Clum (2003, para 2) outline the significant genesis of 'a new assessment approach for higher education … the Collegiate Learning Assessment (CLA) project.' (para., 2). They note that the CLA is 'an assessment that focuses on "general education skills" … [which] measures

students' demonstrated ability to use information.' (para., 10). They suggest that such general education skills include, but are not limited to, critical thinking, analytical reasoning, and written communication. (para., 5). The CLA uses a pre/post testing model that also allows for 'value-added' comparisons within and between institutions (although these are produced differently to the UK version as outlined above). The method of pre/post testing is not a simple individual student T1 to T2 comparison and it is therefore not simply a basic measure of an individual's travelled journey. Rogers (2016) stresses however that over 200 institutions will be using CLA+ (para. 2).

What is also interesting is that the CLA is voluntary but cost $35 in 2016. Students,

it is argued, will likely take the assessment to demonstrate their learning during their time at college ('to justify the price tag.)' (2016, para 1). This may be framed as the higher education 'market' spawning another 'market' to demonstrate that the HE 'product' is worth the cost. This development of a market for testing individual university learning (i.e. the CLA+) is also a change from original institution level data.5 Apparently, a degree result offers insufficient evidence of learning gain, with the CLA+ providing a defining measure of 'learning:'

> students can no longer rely on the collection and mastery of discipline-based information…they need to be able to analyse and evaluate information, solve problems, and communicate effectively to a variety of audiences.6

In an Association of American Colleges and Universities (AAC&U) commissioned study (Hart Research, 2015) of 613 students and 400 employers, the students (as a percentage of the number responding) consistently reported themselves as being 'well prepared' for the workplace. For example, for the skill of 'written communication,' 65% of students responded that they were 'well prepared' as compared with 27% of employers. For critical/analytical thinking, the percentages were 66% of students and 26% of students. The smallest gap was 46% of students and 36% of employers when considering the issue of being 'well prepared' for 'staying current on technologies.' The report also asked employers about 'specific skills for a job' as compared with having gained a 'range of knowledge.' Unsurprisingly, 'both' was the preferred option. However, Arum and Roksa (2010) reported that many students (45%+) do not show any gain in the CLA, in their first two years of HE study and that the figure for the four years of study was 36%. The following section looks to the implications for higher education providers.

5Indeed, one student guide to the CLA + suggests that (for $89) students 'will be able to demonstrate to employers how your 21st century skills attainment compares to students nationwide.'

Council for Aid to Education (CAE) (2012). Student Guide to CLA +.(https://s3.amazonaws.com/StraighterLine/Docs/studentguide_straighterline.pdf (accessed 10.10.18).

6Ibid, p. 1.

3.5 Conclusion

Adopting any measurement of learning gain raises questions about the very nature

and purpose of higher education. If learning gain is defined as positive change in (for example) decision-making skills how might this alter the way in which a degree is taught and/or the actual curricular content of that degree? What pedagogies might be adopted and promoted to facilitate such learning gain(s)? As Biesta (2015) noted, educational purpose must provide a starting point for institutions thinking about measuring learning gain. Figure 3.3 represents the proportion of a university's mission that could be attributed to each of Biesta's categories: qualification (the ability to do something); socialisation (professional values and practices that a student is acculturated to); and subjectification (developing autonomy in thinking and doing). These functions overlap. However, as a trigger for further discussion about the potential influence of learning gain on university mission, we present them here as discrete concepts. Thus, 'Triangle' University sees its mission as 20% subjectification, 50% qualification and 30% socialisation. It is perhaps a liberal arts focused institution and its neighbour, 'Circle University' has a mission of 10% subjectification, 80% qualification and 10% socialisation and is a much more employability/employment focused institution.

Fig. 3.3 A representation of how a university's mission might be proportionally represented

Table 3.2 Analysis of mission statements

Biesta category of purpose	Number of statements
Qualification	4
Subjectification	6
Socialisation	1
Qualification + subjectification	1
Qualification + socialisation	0
Subjectification + socialisation	2
Qualification + subjectification + socialisation	4
No evident educational purpose	2

Earlier analysis (Gossman, Powell, & Neame, 2018), based on an opportunity sample of 21 UK HEIs, found that it was possible to allocate HE mission statements

to the different categories or combinations of Biesta's categories (see Table 3.2). However, the extent to which institutions take seriously the operationalisation of their mission statements in relation to learning of the student experience is debateable. Are institutions intentionally seeking to address multiple purposes here or are these examples, perhaps, of poor institutional focus? Might it be that a learning gain-focused agenda, a purpose reinforced by league tables, pushes universities, and indeed their 'customers,' further towards a qualification-focus? Arguably, for many students, the market is already primarily about qualifications for employment. We invite the reader to find and analyse their own institution's mission statement and to firstly identify it in terms of the Table 3.2 in relation to statements around learning and teaching and then (where there appear to be differing purposes), to try and locate it on Fig. 3.3.

Institutional policies and processes (from marketing for prospective students, to investment choices, for example in relation to student support services) reflect an institution's view of its mission and the overall experience that a student will receive. Arguably, the most important decisions as far as teaching and learning are concerned are not taken at the institutional level. Rather, they are made at programme levels, in respect of the unit of learning. The relationships between learner and teacher are clearly at the heart of the student experience and are where Biesta (2015) categories of purpose will be found, and where the connection between an institution's stated values and purpose are to be given real meaning. As Gibbs (2010, p. 2) concludes:

…what best predicts educational gain is measures of educational process: in other words, what institutions do with their resources to make the most of the students they have.

This points squarely at the notion of learning gain as being something that the best run institutions can make a difference with, through the quality of teacher interactions, resulting in better outcomes for students, not just in terms of learning gain but also in relation to value added. When conversations between teaching academics touch on student motivations, it is common for the view to be expressed however that all that students care about is their grades: pedagogies that attempt to engage them via

more creative teaching approaches are sometimes met with less than full enthusiasm.

It could be argued that approaches such as these would do the best promote the characteristics of Biesta's subjectification category. Teachers' efforts to promote workplace visits and engagement with the subject or discipline outside of the assessed curriculum may not be welcomed by all students, but they may be very effective in the subjectification purpose of the curriculum. There is a tension here between what the purpose of higher education is, the institutional imperative to perform well in league tables, the values and motivations of teachers, and the diversity of learners with different motivations and personal desires. These forces interact in complex ways and will often be pulling in competing directions.

Once beyond a superficial consideration of learning gain and value-added, it becomes apparent that identifying appropriate metrics is more challenging than might be first anticipated. The attraction of objective, empirical measures of performance is clear to see, from the point of view of some stakeholder groups. Whatever the precise funding mechanism for higher education, whether it includes loans systems or not, the stakes are high for national governments, not least in ensuring that money is being well spent and that the system is performing in the national interests. However, when we dig beneath the surface numerous problems can be identified, starting with the heterogeneous nature of higher education institutions and their declared educational missions or purposes, not to mention the desired changes that institutions are seeking to bring about in their students and what they would logically be seeking to measure. This is a criticism that can be made of the HEFCE learning gain projects in that they started from a position of what could be measured without first establishing an adequate understanding of the different purposes of higher education within a wider educational context. Moreover, when we drill down to a more granular level the potential for diversion from an institution's stated purpose is great: discipline and subject characteristics come to bear, as do the values and priorities of teaching academics. In other words, context is important.

An obvious note of caution over learning gain needs to be sounded. If measures of learning gain are based upon what it is practical to measure, then the possibility of unintended consequences becomes very real. For example, there will be pressure on institutions to 'game' the measurements in a similar way as they perhaps do now with other league tables. The impact of linking learning gain to the student taught

experience is hard to predict: so too are the potentially detrimental constraints that might be applied in any quest to improve the measuring of learning gain. A quote generally attributed to Peter Drucker, is apposite here: '[w]hat gets measured gets managed—even when it's pointless to measure and manage it, and even if it harms the purpose of the organisation to do so.'7 It is not difficult to apply such an analysis to the measuring of learning gain. However, in opening up the debate about how best to define the concept of learning gain, we aim to provoke further discussion and exploration of what is meant by the notion of the benefits flowing from higher education, both in terms of our learners, and indeed in respect of wider society.

7See further https://athinkingperson.com/2012/12/02/who-said-what-gets-measured-getsmanaged/ (accessed 30.11.18).

References

Arum, R., & Roska, J. (2010). Academically adrift: Limited learning on college campuses. Chicago, Ill: University of Chicago Press.

Benjamin, R., & Clum, M. (2003). A new field of dreams: The collegiate learning assessment project. Peer review, 5(4), 26–29.

Biesta, G. (2015). What is education for? On good education, teacher judgement, and educational professionalism. European Journal of Education, 50(1), 75–87.

Blackmore, P., & Kandiko, C. (2012). Change: processes and resources. In P. Blackmore & C. Kandiko (Eds.), Strategic Curriculum Change, Global Trends in Universities. Abingdon: Routledge.

Boud, D. (2018). Assessment could demonstrate learning gains, but what is required for it to do so? Higher Education Pedagogies, 3(1), 54–56.

Cameron, A., Wharton, Y., & Scally, J. B. (2018). An investigation into the comparative learning gain and 'value added' for students from widening participation and non-widening participation groups: A case study from sports degrees. Higher Education Pedagogies, 3(1), 83–102.

CEM. (2017). Alis. Retrieved from https://www.cem.org/alis.

Council for Aid to Education (CAE). (2012). Student guide to CLA+. Retrieved from https://s3.amazonaws.com/StraighterLine/Docs/studentguide_straighterline.pdf.

Clarke, M. (2017). Rethinking graduate employability: The role of capital, individual attributes and context. Studies in Higher Education.

Gibbs, G. (2010). Dimensions of quality. International Journal. Retrieved from http://www.heacademy.ac.uk/assets/York/documents/ourwork/evidence_informed_practice/Dimensions_of_Quality.pdf.

Gossman, P., Powell, S., & Neame, C. (2018). Pain, gain—Mission. Higher Education Pedagogies, 3(1), 7–9.

Hart Research Associates. (2015). Falling short: College learning and career success. Washington, D.C.: Hart Research Associates.

Hefce. (2015). Higher Education Funding Council for England. Retrieved from: http://www.hefce.ac.uk/news/newsarchive/2015/Name,105306,en.html.

Jääskelä, P., Nykänen, S., & Tynjälä, P. (2016). Models for the development of generic skills in Finnish higher education. Journal of Further and Higher Education, 1(13), 1–13.

McGrath, C. H., Guerin, B., Harte, E., Frearson, M., & Manville, C. (2015). Learning gain in higher education. Cambridge, UK: RAND Corporation. Retrieved from: https://www.rand.org/pubs/research_reports/RR996.html.

OECD. (2013). Assessment of higher education learning outcomes, feasibility study report. Volume 1 – Design and implementation, Executive Summary. Retrieved from: http://www.oecd.org/education/skills-beyond-school/AHELO%20FS%20Report%20Volume%201%20Executive%20Summary.pdf.

Rogers, K. (2016). A new final exam for college seniors. Fox Business. Retrieved from: http://www.foxbusiness.com/features/a-new-final-exam-for-college-seniors.

Sands, D., Parker, M., Hedgeland, H., Jordan, S., & Galloway, R. (2018). Using concept inventories to measure understanding. Higher Education Pedagogies, 3(1), 173–182. https://doi.org/10.1080/23752696.2018.1433546.

Shavelson, R. J., & Huang, L. (2003). Responding responsibly to the frenzy to assess learning in higher education. Change, 35(1), 11–18.

Sin, C., & Neave, G. (2014). Employability deconstructed: Perceptions of Bologna Stakeholders. Studies in Higher Education, 41(8), 1–16.

The Quality Assurance Agency for Higher Education. (2013). Skills for employability. Retrieved from: http://www.qaa.ac.uk/assuring-standards-and-quality/skills-for-employability.

Thompson, D. (2018). Value added measures in performance tables: A recap of the main issues for primary schools—FFT Education Datalab. Retrieved from: https://ffteducationdatalab.org.uk/

2018/05/value-added-measures-in-performance-tables-a-recap-of-the-main-issues-for-primaryschools/.

Wheadon, C. (2013). Where's the value in 'value-added'? Retrieved from: https://cerp.aqa.org.uk/

blog/bloggers/christopher-wheadon.

Peter Gossman (Ed.D.) is a Principal Lecturer in Education at the University of Worcester, England, within The School of Education and The Department of Education and Inclusion where he

is the programme leader for the post-graduate certificate in learning and teaching in Higher Education. Peter has published on a range of subjects in HE including the scholarship of teaching and

learning.

Stephen Powell (Ph.D.) is a Principal Lecturer at Manchester Metropolitan University, England,

having worked in education for over 20 years, initially as a teacher in the compulsory school sector, and then in HE as a developer of innovative online programmes. He has particular experience in curriculum design and development to meet the needs of learners in the workplace and the

use of inquiry-based approaches to learning, and patchwork text as a form of assessment. He has

developed and managed numerous projects in HE working with colleagues to develop new taught

provision and improve institutions educational systems and processes using action research and

systems thinking. He leads the PGCert in Learning and Teaching in HE as well as running externally funded projects.

C