

The Times Higher Education Supplement

September 13, 2018

Learning gain: political expedient or meaningful measurement?

**SECTION:** NEWS No. 2374

**LENGTH:** 2339 words

The political craving for simple measures of learning gain is neither pedagogically informed nor sufficiently nuanced. Alex Forsythe, Carol Evans, Camille Kandiko Howson and Corony Edwards argue that only by changing focus will the concept become useful

A few years ago, a major UK airport determined to speed up passenger movement by incentivising baggage handlers to move luggage from planes faster. Since it was deemed unfair to compare the overall unloading times of, say, a giant Airbus 380 with a mid-sized Boeing 737, managers agreed to track the time it took for the first piece of luggage to hit the reclaim carousel. The result? An athletic member of the handling team would grab a medium-sized bag from the hold and sprint to the carousel. Everyone else's bags were offloaded at leisure.

Measurement promises to provide transparency and comparability, motivate staff, improve quality and take you out for a beer on a Friday. But the above story, recounted in the 2015

[book](#)

Measurement Madness: Recognizing and Avoiding the Pitfalls of Performance Measurement, is just one of countless examples of measurement's tendency to change behaviour in perverse and cynical ways, rendering all those promises empty.

This is as true in education as in any other walk of life. Alfred Binet, creator of the first IQ tests, imagined that they would improve society by identifying underperforming students in 19th-century France. An advocate of education for all children, the psychologist hoped that his scale comparing mental age with actual age would help teachers better understand pupils' individual needs. However, Henry Goddard, the prominent American psychologist and eugenicist, discovered Binet's work and popularised it with the intention of "curtailing the reproduction of feeble-mindedness". Measurement is never simply a process of identifying objective facts; people decide what to measure, how to go about counting it and what to do with that information.

Measurement, of course, always had its critics, not least in academia. Psychiatrist Leopold Szondi claimed in the 1950s that "the cancer of testology and testomania" had almost completely suppressed actual learning. Yet its advocates continue to invest great hope in it. Now, almost 70 years later, many believe that university education will be transformed - and its value robustly demonstrated - by measures of what is referred to as "learning gain".

There is no one definition of what learning gain encompasses, but most conceptions emphasise changes in students' knowledge and skills, work-readiness and employability or personal development during their time at university. For instance, the Council for Aid to Education's Collegiate Learning Assessment Plus (CLA+) test, widely used to assess learning gain in the US, is an open-ended, question-based test to explore thinking skills, problem-solving, analytic reasoning and communication skills. It typically requires students to evaluate hypothetical scenarios based on real-world issues. A current exemplar on the CLA website requires students to imagine that they are working for an independent organisation tasked with evaluating the claims made by fictitious mayoral candidates, and to make recommendations on which candidate to endorse accordingly.

The concept of learning gain has been widely applied in the US, and variably across the globe over the past couple of decades. Interest from the UK is relatively new, but, in England, it has quickly acquired political favour in the post-2012 era of £9,000+ tuition fees. To its advocates, it appears to provide part of the answer to government concerns around quality control within higher education, as well as value for money in terms of graduate earnings and social justice issues around the lower average outcomes of students from deprived backgrounds.

But while changes in certain student attributes may be measurable, determining cause and effect is more contentious given the diversity of students' lives, both in and out of the university classroom. Moreover, there is a big question about whether learning gain measures should be subject specific. Science, technology, engineering and mathematics subjects, in particular, sometimes take this approach, employing tests that usually comprise multiple choice questions designed to assess subject knowledge. But their reliability and validity is varied, and they are of limited use in explaining the rationale underpinning students' thinking.

This range of approaches is reflected in the

### [13 longitudinal pilot projects](#)

, involving more than 70 higher education institutions, that the Higher Education Funding Council for England - now replaced by the Office for Students - has been funding since 2015 (some of which have now concluded), as well as in the recently cancelled National Mixed Methodology Learning Gain Project, which involved 10 UK institutions and had started to explore methods of measuring learning gain that might be scalable at a national level. Some of the approaches these projects investigate are

discipline-specific. Others focus on more generic topics such as employability or cognitive gains. Some emphasise specific attributes, such as students' academic writing ability or their confidence in their ability to do well.

With taxpayers paying billions of pounds to support UK universities each year, there certainly are some very strong accountability reasons why we should be investing in learning gain, in an attempt to demonstrate the "value added" by a university education. However, this new experiment also risks adding to the data overload and the pressure to perform that many students and academics already experience. Worse, it may add to student anxieties about their own learning journeys. No more will the road less travelled be valued; rather, the question could become: did the student make the effort to travel far and fast enough?

There are other dangers, too: the incentives and reward structures in academia are a study in how not to encourage genuine teaching excellence, with existing performance measures already driving perverse teaching

#### behaviours

, such as teaching to the test and competing for the fastest feedback turn-around times, at the expense of quality and rigorous standards. Indeed, when proposals on learning gain were first floated in England, initial reactions focused on how to game the metrics: savvy institutions, it was claimed, would tell students to dumb down their answers on their entry test, to allow more room for "gain" by the end.

There were also concerns that elite institutions would be unfairly disadvantaged because their students were already so clever (or well coached) when they started. This reaction indicates a fear that some universities may not be doing much more than recruiting the best and the brightest and churning them out a few years later without adding much value along the way.

Other observers worry that data collected about students' experiences are already used well beyond their intended purposes, leading to a plethora of metrics that are not aligned with people and processes. Module-level evaluation data are used to fire individual teaching staff - even if the issues students are not happy with are outside their control. In the UK, student satisfaction has been the dominant focus in the measurement of teaching over the past decade - even though studies show that it is inversely related to meaningful learning since achieving this is often an unsatisfying, disruptive struggle that doesn't necessarily equate with popularity, at least in the short timescales involved. Acknowledgement of its flaws has led the weighting of the National Student Survey (**NSS**) to be halved in the teaching excellence and student outcomes framework (TEF), but this has not made it

any more valid a measure. The question is whether adding learning gain measures to the mix would improve the quality of teaching assessment or just introduce more red herrings.

If applied in a critical way as an integral part of curriculum design and delivery, learning gain as a concept has huge potential to offer valuable insights into the learning process. However, the quest for a universal measure of learning gain is clearly a futile one given the increasing flexibility in precisely what is taught and how it is taught on modern degree programmes - not to mention the specifics of assessment. Standardised measurement does not enable valid comparisons between institutions because those institutions need to be able to adapt and contextualise learning gain relative to the students they serve.

Investment in measures of precisely targeted, empirically tested learning interventions may help students to optimise their potential, but it does not necessarily follow that those interventions are appropriate for all programmes of study, or for all institutions. In some instances, large-scale measurement may be poorly attuned to contextual idiosyncrasies. Similarly, small-scale measures may be of little relevance outside a specific discipline.

Moreover, we should reflect on the fact that if existing assessment regimes were wholly fit for purpose in the first place, there would be no need for additional learning gain measures. We cannot expect one overarching assessment to capture all we need it to do, but we ought to expect standard assessment to do a better job of capturing improvements in key areas of knowledge, skills and understanding, both discipline-specific and generic.

Even if a reliable method of assessing learning gain were discovered, problems would still persist. For example, the difference between a first- and second-year student may be fairly small, so measuring the various gaps between different cohorts would be immensely difficult. As seen in many social science experiments, the spurious noise that researchers pick up on when trying to measure tiny variables often results in false positive and false negative findings. Measures may even track a decline in learning gain - and educators will explain that by arguing that student development is not one unbroken upward ascent towards understanding, and that some undergraduates founder at points in their degree, before picking up the pace.

Tracking a student's learning gain at a specific point in time, such as at the end of a programme of study, also says little or nothing about that individual's learning journey, their progress relative to their peers or their potential for future development. Therefore, while it may be theoretically possible to rank students at the end of a three-year undergraduate degree, it doesn't follow that this will provide a measure of distance travelled that could be used to compare the performance of universities or students - not least because such a measure fails to account for different starting points: a plethora of different entry qualifications, previous educational opportunities and personal circumstances.

If learning gain measurement is genuinely to capture the value-added component of a university experience, attention needs to be directed away from test score improvements towards developing an understanding of the factors contributing to learning. More meaningful measures of learning and teaching, especially in the pursuit of student equality of opportunity, should be considered, even if such approaches may not be generalisable to large populations. This requires rethinking students' transitions into and through higher education, with a clear intent within curriculum design to support students to maximise their social, cultural and political capital: especially important for first-generation or commuter students, as well as those from lower socio-economic backgrounds and with protected characteristics, such as disabilities.

The political craving for simple measures might be a boon for university education departments, providing fruitful avenues for publication, as well as satisfying external, politically driven demands. But researchers must always be explicit about the nature of the measurement process, the validity of learning gain measures and the negative consequences that may follow from adopting them. Student learning and all dimensions of staff development are being undermined by an overzealous emphasis on learning gain that is not pedagogically informed and not attuned to the specific learning context.

All too frequently, learning gain data are used in a reductionist way to gather data from students on their attitudes towards learning, or identifying competencies in specific skills, rather than being attuned to the sensitivities of the taught programme, or contextualised in complex practice. For example, the American CLA+ assumes a degree of standardisation in terms of what is taught and how, but evidence highlights its discipline- and context-dependent nature and the limited reliability of the information it yields for use at the individual level. The relevance of such data to assessments of how students have learned as part of their taught delivery is highly questionable, yet even specialists rarely challenge its plausibility.

In a forthcoming special issue of the journal *Higher Education Pedagogies* dedicated to learning gain, we argue that a far more integrated approach to the concept is needed. Its success will depend upon researchers, practitioners, policymakers and professional services colleagues working together from the outset, knitting together relationships among lecturers' and students' knowledge, concepts and skills, then embedding evidence-driven learning gain approaches within curriculum design and delivery, as well as in accountability measures.

Learning gain cannot merely be a metric-chasing tool presented as a "proof" of quality: it must show what students know, and in what ways, and what works well - when, for whom and why. Only this way will the concept make a meaningful contribution to pedagogy, rather than muddying the waters further, in pursuit of political chimeras. I

Alex Forsythe is a senior lecturer in psychology at the University of Liverpool, Carol Evans is a professor in higher education at the University of Southampton, Camille Kandiko Howson is a senior

lecturer in higher education at King's College London and Corony Edwards is an independent higher education consultant.

**LOAD-DATE:** September 21, 2018