

## Research in Practice

# Measuring student learning gain: a review of transatlantic measurements of assessments in higher education

**Rebecca Randles\* and Alison Cotgrave\*\***

*Faculty of Engineering and Technology, Liverpool John Moores University, Henry Cotton Building, Webster Street, Liverpool, L3 2ET, UK*

Contact: \* [R.Randles@2013.ljmu.ac.uk](mailto:R.Randles@2013.ljmu.ac.uk); \*\* [A.J.Cotgrave@ljmu.ac.uk](mailto:A.J.Cotgrave@ljmu.ac.uk)

### Abstract

*Learning gain is a potentially valuable lens on student development. Its importance has been highlighted by discussions framing the teaching excellence framework (TEF), which considered the 'distance travelled' by a learner as a potential metric. Learning gain in England is at its early stages whilst the United States have had a much more established measurement of learning gain culture. This paper examines the already existing measures being used within the US including the three standardised measures of the Voluntary System of Accountability, for example the Collegiate Learning Assessment, and one self-report measure. A review of these measures illuminates key lessons from existing practice in the US.*

### Keywords

learning gain; learning outcomes; value added; assessment; measures; teaching excellence framework

*Please cite this paper as:*

Randles, R. and Cotgrave, A. (2017) 'Measuring student learning gain: a review of transatlantic measurements of assessments in higher education' in *Innovations in Practice*, 11 (1): 50-59



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 Licence](https://creativecommons.org/licenses/by-nc-nd/4.0/). As an open access journal, articles are free to use, with proper attribution, in educational and other non-commercial settings.

## Introduction

Assessments of student outcomes are common practice in higher education. As stated by the OECD (2008), opinions have differed on how these assessments should be utilised, with some viewing them as an opportunity to identify the best teaching practices and others as an opportunity to compare higher education institutions. Regardless of the reasoning, it is important that the assessments are both valid and reliable in what they aim to measure.

Learning gain is defined as measuring the distance travelled on a student's skills, competencies, content knowledge or personal development (McGrath et al., 2015). In 2015, a study by RAND Europe was commissioned to review the current national and international climate of measures of learning gain (HEFCE, 2017b). Further, with the support of the Department for Business, Innovation and Skills (BIS) and the Higher Education Academy, a report was produced which explored a wide variety of methodological approaches to measuring learning gain and highlighted the need for measures within an English setting (HEFCE, 2017b).

The Higher Education Funding Council for England (HEFCE) began piloting studies into learning gain in 2015, starting with 13 collaborative institutional projects involving over 70 institutions. These projects combined the cross-sectional and longitudinal approaches to learning gain measures as well as a wide range of methodologies (HEFCE, 2016a). Following this, the National Mixed Methodology Learning Gain Project (NMMLGP) was announced. This is a longitudinal research study on learning gain of undergraduate students. 31,000 students were expected to take part in an online assessment programme in each year of their studies to assess their critical thinking, motivation,

attitudes towards their study experience and their engagement (HEFCE, 2016b, HEFCE, 2017a). Complementary to this longitudinal project, HEFCE also announced its Higher Education Learning Gain Analysis (HELGA) programme as a means of comparing known measures by applying them to a large administrative dataset (HEFCE, 2017a). The resources and activities stem from a UK Government Green Paper (BIS, 2015) that identified learning gain as a potential metric in the teaching excellence framework (TEF), offering a possible perspective on teaching quality.

In the US, where learning gain has been more embedded in the culture of higher education, it was brought into sharper focus with the release of *Academically Adrift* (Arum and Roska, 2011). This contained a critical review of the current learning gain measures in the US focusing, in particular, on the Collegiate Learning Assessment (CLA) (Brooks, 2012) though this was later addressed by the Council for Aid to Education (CAE) with the CLA+, (CAE, 2017b). With the development of learning gain at its early stages in England, it is therefore timely to assess what has been achieved in the US and how the measures can be utilised within an English HE context. This paper describes a few measures that exist in the States, which include standardised measures, such as the CLA+ as well as other self-report measures of learning gain.

## Methodology

### *Literature Search*

Corresponding literature searches were conducted in ERIC, Education Research Complete, Academic Search Complete, PsycINFO and PsycARTICLES. These databases were selected owing to the quality

of coverage in educational research, some of which had been previously used in a review of learning gain literature (McGrath et al., 2015). Keywords and strategy were informed by terms that were used in other reviews (McGrath et al., 2015). Keywords included: 'Collegiate Learning Assessment' (CLA, CLA +), 'learning gain', 'value added', 'knowledge gain', 'outcomes', 'skills assessment' (measurement, scale, test), 'standardised tests' (measurement, scale, assessment), 'employability measure' (assessment, scale, test), 'university' (college, higher education, higher education institutions, graduate), 'grades' (marks, score, classification), 'USA' (United States, US, America).

During the review, it became apparent that articles in some journals were more focused and plentiful than others (for example, the *Journal of Educational Measurement and Research in Higher Education* merited closer analysis). There was also the inclusion of 'grey literature' not indexed in those databases, including studies commissioned by educational bodies and agencies (e.g. HEFCE and the OECD).

### *Inclusion-Exclusion Criteria*

For the purposes of this paper, only articles published in the last ten years were chosen. As a methodological consideration, it is worth noting that the terms 'learning gain' and 'value added' are often used interchangeably and it is only upon inspection of the paper as to whether it was valid to be included in this review. Even though these concepts are defined differently, those that fell in line with the

meaning of learning gain were accepted regardless of whether value-added was used.

### *Procedure*

The procedure comprised of three stages:

- **Initial Search;** search of keywords and inspecting articles further if title seemed appropriate
- **Abstract Screening;** screening abstract based on relevance to definition of learning gain
- **Inclusion;** articles were included based upon reading and determining suitability for review

Articles were sorted into the following sections;

- Collegiate Learning Assessment,
- Collegiate Assessment of Academic Proficiency,
- ETS Proficiency Profile,
- limitations of standardised measures, and
- self-reported learning gain.

### *Results*

The Voluntary System of Accountability (VSA) is an initiative in the US which reports on three standardised measures of college learning: ETS Proficiency Profile, Collegiate Learning assessment and Collegiate Assessment of Academic Proficiency, all of which measure slightly different constructs, as shown in *Table 1* (Steedle et al., 2010). According to Liu (2011) these assessments were chosen from sixteen shortlisted instruments as they were deemed the most adequate in measuring improvement in core skills

Assessment	Constructs	Creator
<b>Collegiate Assessment of Academic Proficiency (CAAP)</b> <a href="http://www.act.org/caap">www.act.org/caap</a>	Critical Thinking Writing Mathematics Reading Scientific Reasoning	American College Testing (ACT)
<b>Collegiate Learning Assessment (CLA/CLA+)</b> <a href="http://cae.org/flagship-assessments-cla-cwra/cla/">http://cae.org/flagship-assessments-cla-cwra/cla/</a>	Critical Thinking Analytical Reasoning Problem Solving Writing	Council for Aid to Education (CAE)
<b>ETS Proficiency Profile (EPP) formerly known as Measure of Academic Proficiency and Progress (MAPP)</b> <a href="https://www.ets.org/proficiencyprofile/about">https://www.ets.org/proficiencyprofile/about</a>	Critical Thinking Writing Mathematics Reading	Educational Testing Service (ETS)

Table one: Overview of Assessments, the Constructs They Measure

### *Collegiate Learning Assessment (CLA/CLA+)*

The CLA was developed by the CAE and originally measured several different academic constructs: critical thinking, analytical reasoning, problem solving and writing (CAE, 2017a). The development of the CLA+ led to the measurement of more specific academic constructs (which is included as a ‘performance task’): analysis, problem solving, scientific reasoning, quantitative reasoning, critical reading and evaluation, argument critiquing as well as writing mechanics and effectiveness (CAE, 2017a). The CLA tests students by giving pieces of information that they are asked to assess and make a recommendation based upon them. According to Brooks (2012) this makes the assessment much more relatable as it refers to real problems. The CLA+ added a multiple choice test to the performance task which was intended to allow for the assessment of a wide range of content and skills as well as the ability to compare at a student level (CAE, 2017a).

The Lumina Longitudinal study was a five year long research project, supported by the Lumina Foundation, to investigate the difference between cross-sectional and longitudinal analyses using the CLA as well

as learning gain differences of first year to final year students both within an institution and in comparison to other institutions (Klein, 2009). At the start of the research, 9,167 first year students completed at least one part of the CLA with 1,330 first year students completing all three. 40 per cent of the schools that agreed to take part in the research met the minimum sample size requirements for all three parts of the CLA (Klein, 2009). These schools tested 4,748 at the end of their first year, 2,327 continuing into their second year and 1,675 were followed all the way to their final year (Klein, 2009). Results indicated that cross-sectional and longitudinal scores were consistent, the performance task indicated that students gained as much during their first two years as their last two, however all the analytic writing gains seemed to occur during the final two years (Klein, 2009).

Arum and Roska (2011) conducted a longitudinal study using the CLA and found that students did not make any significant improvement on their critical thinking skills. The authors asserted that this signalled the decline of effective higher education teaching in the US and they suggested that this may be due to research focused faculty, who place their teaching as their second

priority (Lindsay, 2013, Arum and Roska, 2011). However, Oppenheimer et al. (2017) suggested that even though an assessment may not have found any learning gain, this may not be due to a lack of gain but rather that the assessment itself is not sensitive enough to pick up the changes.

The CLA has stimulated debate around its practicality; its scoring is labour-intensive often resulting in a small sample (Possin, 2013). Furthermore, there has been some debate on whether the constructs measured by the CLA are applicable to all subjects. Brooks (2012) argued that the CLA does not test humanities subjects particularly well as students in these subjects are taught to evaluate information in a different way to other subjects.

### *Collegiate Assessment of Academic Proficiency (CAAP)*

The Collegiate Assessment of Academic Proficiency (CAAP) was created by American College Testing (ACT) and measures five academic constructs; critical thinking, writing, mathematics, reading and scientific reasoning (ACT, 2017). An example of an institution that utilises the CAAP is Calhoun Community College (CCC). The College uses the CAAP highlight particular problems and trends enabling its leaders to take the necessary steps towards addressing any concerns. For example, in their 2012 report (CCC, 2012), out of a total of 674 students, 28 per cent were found to be scoring at below the national mean in writing skills; 55 per cent were below average for critical thinking skills. In the 2014 report (CCC, 2014) out of total of 486 students, a different picture emerged; 55 per cent were found to be below average in writing while 68 per cent were below average in critical thinking skills. In 2016, which had a different focus (CCC, 2016), out of a total of 546 students, 48 per

cent were found to be below the national average on reading while as 63 per cent were below the national average in mathematical skills.

The Wabash National Study also utilises a sub-section of the CAAP for its measure of critical thinking (Blaich and Wise, 2011). This section is a 32-item instrument that measures skills in clarifying, analysing, evaluating and extending arguments (Center for Inquiry at Wabash College [CIWC], 2016). This study was a longitudinal assessment project with the aim of providing information to improve student learning and enhance the educational impact of their studies (CIWC, 2016). This research went beyond just using the constructs of the CAAP by utilising a wide variety of scales, such as the student experience survey in order to attempt to achieve the aim of having high-quality data for institutions to promote improvements in student learning (Blaich and Wise, 2011). Since the pilot stage of this research in 2005, over 17,000 students from 49 HE institutions have joined the study, with 30 in the new version of the study in 2010 (CIWC, 2016). Findings of this study have reported that, whilst students grew in areas such as critical thinking and moral reasoning, there was a decline in areas such as academic motivation. The outcomes of this ongoing research have allowed institutions to judge the need to make steps towards helping with student motivation (Blaich and Wise, 2011). The Wabash study has been the focus of HEFCE's NMMLGP.

### *ETS Proficiency Profile (EPP)*

The ETS Proficiency Profile (EPP), formerly known as the Measure of Academic Proficiency and Progress (MAPP), was created by the Educational Testing Service (ETS) and measures four different constructs; reading, critical

thinking, writing and mathematics. It applies 108 multiple choice questions, or 27 questions per construct (ETS, 2017). Liu (2011) conducted a study looking at the differences in performance between first and final year students using the EPP; 4,373 first year and 1,823 final year students participated. Results indicated that there was a difference between first and final year students; first year students' score average was 110 on critical thinking whereas final year students scored 113. Similarly, on the writing aspect of the test, first year students averaged 113 and final year students 115, suggesting that the measurement is able to distinguish between the two levels of students (Liu, 2011). In addition, it was found that student SAT (Scholastic Assessment Test) scores were moderately correlated with performance on the EPP in both of the constructs; in other words, the measure has construct validity.

Roohr et al. (2016) conducted a longitudinal study of 168 students from a large university in the US using the EPP; the university in question had a retention rate of over 90 per cent from first to final year. Students undertook the test when they first entered the university and they were tested again later in their degree. The time between the two testing periods ranged from four to 55 months; most students took the second test after more than one year (one third of the students were taking STEM subjects as their chosen degree). For the total EPP scores, results became more significant the longer the time between testing, for example when the second test was after three years  $p < .01$ , whereas after four or five years  $p < .001$ . This was also found to be the case on the reading construct and mathematics, however in critical thinking and writing the difference was only found to be significant after retesting after a four or five-year gap.

However, limitations to the EPP have been noted. Lakin et al. (2012) found that those whose first language was English significantly outperformed those whose first language was not English on all domains of the EPP. This indicates difficulties in the wording of questions to some students. However, it was found that the gap was much smaller in the Mathematics domain.

### *Limitations of standardised measures*

When considering the assessment measures there have been papers that offer a broader sense of the tools, particularly in terms of their appropriateness, applicability and practicality. The University of California, for example, rejected the use of standardised measures to assess student learning gain as it was believed that the tests failed to recognise the diversity, breadth and depth of discipline specific knowledge and learning (Douglass et al., 2012). In addition, Banta and Pike (2012) argued that the skills and outcomes that are measured by standardised testing is only a small minority of the learning gain that is taken from completing a higher education degree.

Considering the Wabash National Study, although assessments were a key part of this research, the issue was not the learning gain measures themselves, but rather the difficulty that institutions have in identifying and implementing changes (Blaich and Wise, 2011; Kuh et al., 2014). This highlights a problem with the assessments themselves. Although they are able to measure the learning gain outcomes of students, they fail to identify the reasons why a student is at a certain level as well as failing to suggest changes that the institution can put in place to improve student performance (Klein et al., 2007).

### *Self-reported learning gain*

Self-reports are another method which has been used in academic institutions and involves asking students to report on their own skills (Bowman, 2010). The Student Experience in the Research University Survey (SERU-S) is the only self-assessment survey that is part of the VSA initiative. The SERU-S asks students to rate their level of proficiency at two different time points on a six-point scale on a variety of educational constructs such as critical thinking, writing, reading and comprehension skills (Douglass et al., 2012). Douglass et al. (2012) examined the differences on each of the constructs from the beginning of their university course and compared this to the students' grade point average (GPA). It was found that students reported less learning gain in quantitative skills, which made sense in relation to the students' chosen subject. The biggest gain was found to be in subject knowledge. Higher GPA was also found to be connected to higher levels of self-reported learning gain, suggesting students are able to accurately report their performance over time. In support of this, and in terms of demographics, those who were not born in the US reported the higher learning gain in oral communication. Owing to this apparent veracity, Douglass et al. (2012) propose using self-reports on a more regular basis.

Self-reports can also be highly valuable in improving teaching quality and student academic success. The College Success Factors Index (CSFI) is a self-reporting instrument that has recently been introduced in the US and focuses on ten student success factors grounded in research (Cox and Lemon, 2016). Students are tested at the start and end of the academic year to measure any gain. This assessment looks at areas such as task planning, time management and engagement. A study by Cox and Lemon (2016) found that using interventions to improve the teaching of

these academic gains, improved students' performance across two years and assert that self-report assessments are effective in measuring the changes in teaching quality and should not be completely disregarded.

However, there are limitations to the use of self-reports. The validity of these measures has been widely discussed. According to Bowman (2010) and Porter (2013), the problem occurs as students may report that a certain construct has improved when it has not, or they may report no improvement when they really have. Supporting this, many self-report methods choose to apply multiple-choice as their method of testing owing to its objective scoring and cost effectiveness. Yet, many researchers such as Nicol (2007) have argued that these kinds of tests do not encourage students to engage in higher order processing, owing to ease of selection, which may influence the results. It is also possible that a student could have selected the correct answer without having an ability to understand the question (Nicol, 2007).

Hyytinen et al. (2015) compared the CLA with a multiple-choice test, which both claimed to measure the same constructs but were assessed using different methods; third of students performed better on the multiple-choice test while a quarter performed better on the CLA. This research also the idea that multiple-choice tests simply do not engage a high level of processing.

Bowman (2010) conducted a study that compared subjective longitudinal assessments with students' self-reports of their learning gain. Human errors in judgement were thought to be the biggest hurdle in self-reports although, it is difficult to say for certain whether these errors were simply due to the construction of the self-report. In addition, it was also found that if self-reports are to be utilised then they

should be applied on a longitudinal basis rather than ask students to reflect on their development.

weaknesses of teaching excellence in the sector?

- *Becky Randles is in the first year of her PhD study at the School of Built Environment.*

### Conclusions

Only 28 per cent of community colleges use one of the VSA standardised assessment measures. Although this figure has increased in recent years, it may suggest that the time-consuming (qualitative) elements are unappealing (Liu and Roohr, 2013). There is still much debate on the appropriateness of the CLA, particularly as it has been viewed to lack constructs that are generalised to all degree subjects (Pike, 2015, Brooks, 2012). The MAPP's problems are associated with the nature of multiple-choice testing (Liu, 2011, Hyytinen et al., 2015) whilst the CAAP is thought to be too time-consuming. Self-report methods are more cost-effective and easier to administer, but they come with their own set of problems. Human error, or the accuracy with which students view their performance and learning gain over time are thought to be its major pitfalls (Bowman, 2010). Douglass et al. (2012) study on the link between self-reported learning gain and GPA, demonstrates that grades should not be ignored as a measure of learning gain.

With the evidence presented in this paper, it is difficult to conclude the best way to proceed with learning gain measures in English HE. Each of the different methods seem to have their own pros and cons and it may be impossible to derive a measure that is objective, effective and able to engage students. In terms of the debate within England, key questions remain. Who, and for what purpose, are we measuring learning gain? If it is for the student, how can the data be used to support student development? Finally, and crucially to current debates, how can learning gain metrics be used to identify the strengths and



## References

- American College Testing (2017) 'ACT CAAP', retrieved from: [www.act.org/caap](http://www.act.org/caap) (accessed May 2017)
- Arum, R. and Roska, J. (2011) *Academically Adrift: Limited Learning on College Campuses*, Chicago: University of Chicago Press
- Banta, T.W. and Pike, G.R. (2012) 'Making the case against - one more time...', National Institute for Learning Outcomes Assessment, retrieved from: <http://www.learningoutcomeassessment.org/documents/HerringPaperFINAL.pdf> (accessed June 2017)
- BIS (2015) *Green Paper - Fulfilling our Potential: Teaching Excellence, Social Mobility and Student Choice*, London: Department for Business Innovation and Skills.
- Blaich, C.F. and Wise, K.S. (2011) *From Gathering to Using Assessment Results: Lessons from the Wabash National Study*, Urbana, IL: National Institute for Learning Outcomes Assessment
- Bowman, N.A. (2010) 'Can 1st-year college students accurately report their learning and development?' In *American Educational Research Journal*, 47 (2): 466-96
- Brooks, P. (2012) 'Outcomes, testing, learning: what's at stake?' In *Social Research*, 79: 601-11
- Council for Aid to Education (2017a) 'CLA + for Higher Education', retrieved from: <http://cae.org/flagship-assessments-cla-cwra/cla/> (accessed May 2017)
- Council for Aid to Education (2017b) 'Comparing CLA to CLA+' retrieved from: [http://cae.org/images/uploads/pdf/cla\\_vs\\_cla\\_plus.pdf](http://cae.org/images/uploads/pdf/cla_vs_cla_plus.pdf) (accessed May 2017)
- Calhoun Community College (2012) *CAAP 2012: Collegiate Assessment of Academic Proficiency*, Calhoun Community College
- Calhoun Community College (2014) *CAAP 2014: Collegiate Assessment of Academic Proficiency*, Calhoun Community College
- Calhoun Community College (2016) *CAAP 2016: Collegiate Assessment of Academic Proficiency*, Calhoun Community College
- Center for Inquiry at Wabash College (2016) *Wabash National Study 2006–2012*, retrieved from: <http://www.liberalarts.wabash.edu/study-overview/> (accessed May 2012)
- Cox, T.D. and Lemon, M.A. (2016) 'A curricular intervention for teaching and learning: measurement of gains of first-year college student learning' in *Journal of the Scholarship of Teaching & Learning*, 16: 1-11
- Douglass, J., Thomson, G. and Zhao, C.M. (2012) 'The learning outcomes race: the value of self-reported gains in large research universities' in *Higher Education*, 63(3): 317-35
- Educational Testing Service (2017) 'ETS Proficiency Profile' retrieved from: <https://www.ets.org/proficiencyprofile/about> (May 2017)
- HEFCE (2016a) 'Learning Gain', retrieved from: <http://www.hefce.ac.uk/lt/lg/> (accessed April 2017)
- HEFCE (2016b) 'National Mixed Methodology Learning Gain Project' retrieved from: <http://www.hefce.ac.uk/lt/lg/mixed/> (accessed April 2017)
- HEFCE (2017a) 'HELGA – Higher Education Learning Gain Analysis' presentation at the HEFCE Learning Gain (Pilot Projects) Meeting, 26 April, London
- HEFCE (2017b) 'Learning gain in higher education' retrieved from: <http://www.hefce.ac.uk/pubs/rereports/year/2015/learninggain/> (accessed May 2017)

- Hyytinen, H., Nissinen, K., Ursin, J., Toom, A. and Lindblom-Ylänne, S. (2015) 'Problematising the equivalence of the test results of performance-based critical thinking tests for undergraduate students' in *Studies in Educational Evaluation*, 44, 1-8
- Klein, S. (2009) *The Lumina Longitudinal Study: Summary of Procedures and Findings*, retrieved from: [http://cae.org/images/uploads/pdf/12\\_CLA\\_Lumina\\_Longitudinal\\_Study\\_Summary\\_Findings.pdf](http://cae.org/images/uploads/pdf/12_CLA_Lumina_Longitudinal_Study_Summary_Findings.pdf) (accessed March 2017)
- Klein, S., Benjamin, R., Shavelson, R. and Bolus, R. (2007) 'The collegiate learning assessment: facts and fantasies' in *Evaluation Review*, 31: 415-39
- Kuh, G.D., Jankowski, N., Ikenberry, S.O. and Kinzie, J. (2014) 'Knowing what students know and can do: the current state of student learning outcomes assessment in US colleges and universities' retrieved from: <http://www.learningoutcomesassessment.org/documents/2013%20Abridged%20Survey%20Report%20Final.pdf> (accessed June 2017)
- Lakin, J.M., Elliott, D.C. and Liu, O.L. (2012) 'Investigating ESL students' performance on outcomes assessments in higher education' in *Educational and Psychological Measurement*, 72: 734-53
- Lindsay, T. (2013) 'The likelihood of higher education reform' in *Society*, 50: 236-44
- Liu, O.L. (2011) 'Measuring value-added in higher education: conditions and caveats – results from using the Measure of Academic Proficiency and Progress (MAPPTM)' in *Assessment and Evaluation in Higher Education*, 36: 81-94
- McGrath, C. H., Guerin, B., Harte, E., Frearson, M. and Manville, C. (2015) *Learning Gain in Higher Education*, Cambridge: RAND Europe
- Nicol, D. (2007) 'E-assessment by design: using multiple-choice tests to good effect' in *Journal of Further and Higher Education*, 31 (1): 53-64
- OECD (2008) *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-added of Schools*, OECD Publishing, retrieved from: <http://www.oecd-ilibrary.org/docserver/download/9608101e.pdf?expires=1497189511&id=id&accname=guest&checksum=9AAE9E2C23DEF624E0FFB9E32E7A7C9A>
- Oppenheimer, D., Zaromb, F., Pomerantz, J.R., Williams, J.C. and Park, Y.S. (2017) 'Improvement of writing skills during college: a multi-year cross-sectional and longitudinal study of undergraduate writing performance' in *Assessing Writing*, 32: 12-27
- Porter, S.R. (2013) 'Self-reported learning gains: a theory and test of college student survey response' in *Research in Higher Education*, 54: 201-26
- Possin, K. (2013) 'A fatal flaw in the Collegiate Learning Assessment test' in *Assessment Update*, 25: 8-12
- Roohr, K.C., Liu, H. and Liu, O.L. (2016) 'Investigating student learning gains in college: a longitudinal study' in *Studies in Higher Education*, retrieved from: <http://www.tandfonline.com/doi/pdf/10.1080/03075079.2016.1143925?needAccess=true>
- Steedle, J., Kugelmass, H. & Nemeth, A. (2010) 'What do they measure? Comparing three learning outcomes assessments' in *Change*, 42: 33-37