Piloting the CLA in Ontario

Mary Catharine Lennon
Higher Education Quality Council of Ontario

Cite this publication in the following format:

Lennon, M. C. (2014). *Piloting the CLA in Ontario.* Toronto: Higher Education Quality Council of Ontario.

Table of Contents

Introduction

In 2011, as part of a comprehensive research agenda on learning outcomes development and measurement, the Higher Education Quality Council of Ontario (HEQCO) began supporting eight Ontario institutions to assess the generic skills acquisition of their students. This report summarizes the activities and results of the eight institutions that piloted the Council for Aid to Education's Collegiate Learning Assessment (CLA), a written examination designed to assess the critical thinking and problem solving skills of entering and graduating students. It reviews the rationale for the project, the challenges and issues encountered with CLA test administration and implementation, and the institutions' impressions of the value of the resulting data. While there is significant interest from institutions and programs in measuring the generic skills of students and understanding the amount of learning that can be attributed to the institution, the experiences of the institutions that participated in this project highlight certain administrative and methodological challenges that arise in the move from theory to practice in large scale assessments.

Background

A population that is creative and has strong critical thinking and problem solving skills is the cornerstone of a successful knowledge-based economy (Drucker, 1966; Florida, 2002; Rooney, 2003; Martin & Florida, 2009). Providing students with these skills is imperative to develop a strong workforce and is a primary task of postsecondary education (World Bank, 2010; OECD, 1996). To ensure the competitiveness of future generations, governments, institutions and researchers alike have been seeking to understand and assess students' skills development in these areas for more than a decade (Tremblay, Lalancette & Roseveare, 2012). Making these skills sets more transparent helps an education system identify successes in providing quality education, students articulate their capacities, and allows employers to understand the abilities of recent graduates. This work on developing, articulating and demonstrating the abilities of graduates is critical as it provides insight into the value and contribution of postsecondary education to both individuals and society at large – topics frequently debated in today's society (Hart McKiernan, 2013; Benjamin, 2013).

As part of a broad research agenda to understand and demonstrate the value of postsecondary education, HEQCO undertook a number of research projects to explore learning outcomes – what a postsecondary student knows and can do upon graduation. Since 2011, HEQCO has been engaged in a multitude of activities in the area, including establishing learning outcomes across sectors of disciplines (the Tuning project; see Lennon, Frank, Humphreys, Lenton, Madsen, Omri & Turner, forthcoming); measuring learning outcomes through participation in an international assessment of discipline-level learning outcomes (the AHELO Feasibility Study; see Lennon & Jonker, 2014); as well as supporting the incorporation and demonstration of learning outcomes through a variety of institutional activities.

While the majority of HEQCO's activities in the area of learning outcomes focus on sectors, disciplines or institutional activities, it is clear that there are also generic skills – for example, verbal and quantitative reasoning, comprehension, critical thinking and problem solving, communication – that transcend these boundaries and are cultivated regardless of content-specific knowledge or institutional priority/focus.[1] These generic skills constitute the primary competencies that students expect to gain, the traits that employers seek, and the basis for a strong society and economy (Florida, 2002; Martin & Florida, 2009; Tremblay, Lalancette & Roseveare, 2012).

---

[1] In fact, the HEQCO Tuning project determined that the majority of generic competencies were common across disciplinary sectors (Lennon et al., forthcoming).

Value of Generic Skills Performance Assessments

Decades of popular commentary suggest that a student chooses to attend postsecondary education – particularly university – in order to learn to "think". Upon graduation, a student would be expected to have acquired basic "thinking skills" and know how to apply them to new situations (Pellegrino, Chudowsky & Glaser, 2001). These capacities would in turn ensure success in the labour force. But, in a world where 59% of Ontarians have a postsecondary education (Norrie & Lin, 2009), the assumption that postsecondary graduates have acquired these generic skills – ungrounded in evidence or fact – no longer satisfies the public demand for accountability in demonstrating student learning gains.[2] In the current environment, government, institutions and students have been under pressure to demonstrate the actual skills, competencies and abilities of graduates: in short, to demonstrate with evidence the value of a postsecondary credential.

Many assessments, both large-scale and in the classroom, are successful at gauging students' content knowledge in their discipline or sector.[3] However, fewer are targeted at the broader knowledge expected of students.[4] Some such assessments do exist, such as the Graduate Record Examination (GRE) General Test[5], the Australian Graduate Skills Assessment[6] and the Brazilian ENADE[7], and are used to serve various stakeholder purposes. None of these large-scale assessments claims to measure everything gained from education, but rather highlights certain student abilities (be they generic or discipline-specific).

The primary stakeholders of generic skills tests, commonly accepted to include governments, institutions, employers and students, benefit from assessments of learning outcomes in different ways:

- Governments gain:
  - information on the quality of graduating students
  - performance indicators and accountability measures
  - the ability to track system-wide improvement over time
  - potential for international benchmarking
- Postsecondary institutions gain:
  - identification of at-risk students (when conducted on entry)
  - information on program effectiveness and improvement over time
  - information for curriculum development
  - institutional benchmarks
  - information for the selection of graduate students
- Employers gain:
  - information for employee recruitment
  - evidence of proficiency of skill sets
- Students gain:
  - information on their skills sets and areas for improvement
  - proof of performance to provide to potential employers

---

[2] Certainly there is labour market value to PSE, and the earning differentials and employment rates of those who attend are significantly higher than for those who do not (Tal & Enenajor, 2013).
[3] I.e., the ETS Proficiency Profile (see https://www.ets.org/proficiencyprofile/about/), and professional accreditation exams such as the Engineering Professional Practice Exam (see http://www.peo.on.ca/index.php/ci_id/2060/la_id/1.htm).
[4] The OECD's AHELO Feasibility Study focused on both generic skills and discipline-specific skills.
[5] For more information, see http://www.ets.org/gre.
[6] For more information, see http://www.acer.edu.au/tests/gsa.
[7] The Exame Nacional de Desempenho dos Estudantes (ENADE) tests both generic and discipline-specific skills.

How, when and by whom an assessment is administered determines the prioritization of the benefits to stakeholders. Currently, no one test is suitable to address all stakeholder needs equally. However, the Council for Aid to Education's Collegiate Learning Assessment (CLA) has been described by some as the most comprehensive tool available (Arum & Roksa, 2011, p. 27) and has attracted international attention, with many countries, including Australia and Ireland[8], exploring its use. Furthermore, the CLA was also used as the basis for the Generic Skills Assessment in the OECD's AHELO Feasibility Study (OECD, 2009), making it central to the assessment of learning outcomes on the international stage.

Because generic skills are fundamental to understanding the value of postsecondary education, and because their measurement remains elusive, the Higher Education Quality Council of Ontario was interested in the movement towards large-scale assessments. The prominence of the CLA in the early movement toward the assessment of learning outcomes around the world made it an interesting and attractive tool for the purposes of this project. HEQCO supported eight colleges and universities to pilot the CLA and its college-sector equivalent, the CCLA (Community College Learning Assessment), to consider its value to institutions, and subsequently determine its usefulness and effectiveness for government, employers and students.

The Collegiate Learning Assessment

Launched in the United States in 2002 by the Council for Aid to Education (CAE), the CLA/CCLA is a standardized test that assesses a postsecondary institution's contribution to the development of students' key competencies, including critical thinking, analytic reasoning, problem solving and written communication skills. The 90-minute written test consists of three online assessment components – two analytical writing tasks and a performance task.[9] The tasks are realistic problems that require students to analyze complex materials to create a cogent and well-constructed argument.

A key component of the CLA is that it assesses the "value-add" of education, a term used since the late 1960s in K-12 education to refer to the measure of effectiveness of a school (Office for Education Policy and Innovation, 2007). Rather than relying on one-time scores of performance, the "value-add" model of assessment accounts for external factors, such as socioconomic status or prior attainment (such as high school grades), which are highly correlated to student performance but are not a reflection of a school's contribution to student achievement.[10] The "value-add" is calculated by determining the difference between students' expected performance, projected using test scores corrected to account for external factors, and students' actual performance. This measure can then be directly related to the quality of the educational experience.[11] In the United States, the CLA uses SAT scores as the benchmark of external factors. Institutions are provided with benchmarks on how their students as a whole are performing compared to other institutions, and how much progress their students have made since entry. This provides a better understanding of the institutional contribution to student learning and highlights strengths that can lead to overall improvements in teaching and learning.

---

[8] Neither Australia nor Ireland has subsequently implemented the CLA. See Australian Government (2011) and HETAC (2011) for a description of their activities.
[9] The CLA+ is a modified version of the CLA introduced in fall 2013. As the Ontario pilots engaged with the original assessment tool, all commentary is related specifically and exclusively to the CLA.
[10] A formidable review of "value-add" purposes and techniques is provided in OECD (2013).
[11] In the case of the Ontario pilot sites performance on the Scholastic Level Exam (SLE) was used as a benchmark for student performance.

Ontario's Pilot Projects

In the fall of 2011 HEQCO partnered with eight institutions from across Ontario to pilot the tests in a variety of faculties and programs. HEQCO offered financial and coordinating support, as well as a platform for discussing issues and challenges, but did not have access to student or institutional data. Each institution provided interim and final reports outlining feasibility of administration, value of data, and potential for broader use of the assessment tool. Each institution was independently responsible for the administration of the assessment, data analysis and reporting, and for providing students with feedback.

The partner institutions included:

- Fanshawe College – Schools of Business and Information Technology
- Humber College Institute of Technology and Advanced Learning – Business Administration co-op (3-year), Business Administration (3-year), Business Management (2-year)
- McMaster University – Faculty of Social Sciences
- Mohawk College and McMaster University – Collaborative Medical Radiation Science Program
- Queen's University – Faculty of Engineering and Applied Science
- Sheridan College Institute of Technology and Advanced Learning – Degree programs (not offered in collaboration with a university)
- University of Guelph – Bachelor of Arts Honours, Bachelor of Science Honours
- University of Windsor – Faculty of Arts, Humanities and Social Sciences; Faculty of Engineering; and Faculty of Human Kinetics

Six institutions conducted cross-sectional studies, while the remaining two attempted a combined cross-sectional and longitudinal approach.[12] The longitudinal approach sought to test individual students two or more times over the course of their programs to assess the progress in their learning. The cross-sectional approach tests a sample of first-year students (in the fall) and final-year students (in the spring) and examines the differences in the two populations. The longitudinal approach can be more robust, but the cross-sectional approach yields results more quickly. Neither institution completed its longitudinal component due to low participation rates in one case and perceived lack of value in the other.

Timelines for administering the CLA/CCLA varied depending on program length and methodological approach, but the majority of institutions tested students in the spring and fall of 2012.

Activities and Results

The administration of the test posed a number of challenges, many of them common across programs and institutions, which impacted the success of the pilots. To assess students, each of the institutions required consent from their institutional research ethics board. In many cases this did not present a significant hurdle, but some teams were required to modify their intended administration technique (from mandatory to voluntary), the recruitment techniques used, the inducements/incentives offered, or the linking of study data to administrative data. In one case, resolution of the research ethics clearance resulted in a six-month delay in test implementation, in part owing to the narrowness of the testing windows.

---

[12] One institution designed a longitudinal component but abandoned it after obtaining a low response rate in the first administration. They continued with the cross-sectional approach.

Most institutions sought voluntary participation from students outside of class time. The recruitment process was one of the most time consuming and costly activities of the pilots. Institutions noted the complexities of organizing session times and locations, two critical elements to ensuring that the test was convenient for students in order to maximize the likelihood of participation. Contacting and advertising to students was also a time consuming process. It was understood that failing to achieve an appropriate sample of participants would have significant implications on the reliability of the data, and was thus a primary concern for institutions. This was complicated by the relatively short CAE window during which the test was available to students (between February 1-April 15 and August 15-October 31), which conflicted with some institutional academic calendars.

Some programs had greater ease with recruitment and administration than others, with smaller, more focused programs reporting fewer challenges. An explanation for this phenomenon might be found in the "cohort effect", which suggests that it is easier to motivate smaller groups than large, diverse populations (Tremblay, Lalancette & Roseveare, 2012, p. 162). The University of Guelph, which targeted students in the arts and science degree programs, noted that the culture of certain programs made it more difficult to recruit volunteers for testing. Likewise, Fanshawe College suggested that the lack of buy-in and support from faculty members was a possible reason for low recruitment numbers.

Despite great efforts, many institutions had significantly lower participation rates than expected, which in turn hindered their ability to conduct data analysis (see below for further commentary on data). In the most extreme case, Humber College did not acquire enough participants in its first round of administration to warrant a second administration, nor was it possible to run the longitudinal aspect. Thus, the further assessments were cancelled and the pilot project was abandoned. Similarly, Fanshawe College was unsuccessful in its recruitment from the general population and shifted its efforts to focus explicitly on international students, who made up a very responsive population.

Two pilots chose to embed the assessment into program/class time. While this increased participation rates, it also required a considerable amount of early planning as well as ethics approval at each institution. Queen's was able to embed the assessment for first year-students as part of a larger research project piloting a variety of assessment tools and made participation mandatory.[13] Similarly, at McMaster/Mohawk, students could either participate in the CLA or complete another assignment for 5% of their final course grade.

The Humber project aimed for an embedded census of students in both their first and fourth years and attempted to use the CLA in place of another institution-wide standard assessment of reading and writing (used by faculty members to identify at-risk students who may require remedial language programming). Because the CLA is a proprietary product, student responses are not made available to faculty members or institutions. Hence, the CLA is not a tool that provides early detection information or allows for other uses, and thus the ethics board would not allow it to replace the mandatory test.

Most administration issues were not specifically related to the CLA study but could instead have occurred in any large-scale, low-stakes assessment. CLA-specific elements that impacted the administration were linked primarily to the test window and the strict propriety nature of the exam, which made it impossible to use the CLA for any concurrent purpose (as highlighted in the case of Humber).

The actual data and results are the key component to understanding the perceived success of the CLA as an instrument to measure generic skills. Considerations related to the reliability of the data collected, the ability to link them with other data sets, and the potential use of the information gathered were critical for determining the potential uses and contribution of the CLA.

---

[13] Queen's final-year students were volunteers.

Institutions noted a number of concerns in using the data gathered from the CLA as a source of information. Many of the concerns were based on the small sample sizes, self-selection bias of the voluntary participants, and the uncertainty of student effort on the assessment. These are not issues exclusive to the CLA but could be encountered with any large-scale, low-stakes assessments – many of which could be mitigated through careful research design at the institutions.

To explore the value of the data, institutions with sufficient data indicated that they were able to link the data to other institutional sets, including demographic cohort comparisons, admission averages, score variations, variations among those who volunteered to be tested and those who did not, etc. One institution, for example, found the CLA to be correlated with some data sets (i.e., admission averages and international student status) but not others (i.e., course grades). This provided an indication of the "value-add" in one institution, but was not robust enough for conclusive evidence. Unfortunately, most institutions were unable to conduct any analysis of the data, as the sample sizes were too small.

In considering the potential uses of the CLA data, the few institutions that found the data valuable suggested that the data could potentially – upon further investigation – be used to assess critical thinking and writing, for benchmarking against other institutions, identifying institutional value-add, and to evaluate the effectiveness of certain teaching strategies.

Discussion

The purpose of piloting the CLA was to determine the usefulness of this test as an appropriate mechanism for assessing generic skills in Ontario's postsecondary education system. Institutional responses to this question were mixed, with several common responses emerging across the institutions and programs primarily relating to administrative issues, related issues in data collection methodologies, and concerns with the usefulness of the data as a means to comment on the abilities of broad student populations.

Some of these issues have been addressed in the newly developed CLA+.[14] CLA+, for example, may mitigate issues of student recruitment and reduce concerns about student effort by providing more significant individual level feedback. This provides students with a comparative understanding of their personal abilities, a reflective learning tool (if in first year), and a record of achievement to supplement their credential (if graduating). Making students a central focus of the CLA+ may present this low-stakes assessment as a more attractive activity for students, thus reducing the administrative burden of recruitment – one of the most challenging aspects of the Ontario pilots. Institutions agreed that the assessment of learning outcomes and generic skills was a high priority for postsecondary institutions and programs. Recognizing that there is no "silver bullet" to evaluate generic skills acquisition in postsecondary education easily, the CLA pilots provided insight into how large-scale assessments can contribute to our understanding, and ways in which institutions want to use the information. It also provided considerable insight into issues of administration that should be considered in future research design methodologies.

---

[14] For more information on the CLA+ and how it differs from the CLA, please see: http://cae.org/performance-assessment/category/cla-overview/.

References

Arum, R., & Roksa, J. (2011). *Academically Adrift: Limited Learning on College Campuses.* Chicago and London: The University of Chicago Press.

Australian Government. (2011). *Assessment of Generic Skills: Discussion Paper*. Australia: Australian Government. Retrieved from http://www.innovation.gov.au/HigherEducation/Policy/Documents/AssessGenericSkills_Finaldiscussion paper.pdf

Benjamin, R. (2013). Teach, Learn, Assess. In K. Norrie & M. C. Lennon (eds.), *Measuring the Value of a Postsecondary Education* (pp. 123-140). Montreal and Kingston: McGill-Queen's University Press.

Drucker, P. (1966). *The Effective Executive*. New York: Harper Business.

Florida, R. (2002). *The Rise of the Creative Class*: *And how it's transforming work, leisure, community and everyday life.* New York: Basic Books.

HETAC. (2011). *Assessing the Impact of Higher Education on Learners' Skills: Research report on a pilot study of learners' critical thinking, problem solving and analytic reasoning skills in the HETAC sector*. Dublin, Ireland: HETAC.

Hart McKiernan, H. (2013). Increasing Attainment in the United States: Quality is about learning – What students know, understand and are able to do. In K. Norrie & M. C. Lennon (eds.), *Measuring the Value of a Postsecondary Education* (pp. 39-52). Montreal and Kingston: McGill-Queen's University Press.

Lennon, M. C. (2010). *Signalling Abilities and Achievement: Measuring and reporting on skill and competency development*. Toronto: Higher Education Quality Council of Ontario.

Lennon, M. C., Frank, B., Humphreys, J., Lenton, R., Madsen, K., Omri, A., & Turner, R. (forthcoming). *Tuning: Identifying and Measuring Sector-Based Learning Outcomes in Postsecondary Education*. Toronto: Higher Education Quality Council of Ontario.

Lennon, M. C., & Jonker, L. (2014). *The AHELO Feasibility Study in Ontario*. Toronto: Higher Education Quality Council of Ontario.

Martin R., & Florida, R. (2009). *Ontario in the Creative Age*. Toronto: Martin Prosperity Institute, Joseph L. Rotman School of Management, University of Toronto.

Norrie, K., & Lin, S. (2009). *Postsecondary Educational Attainment and Participation in Ontario*. Toronto: Higher Education Quality Council of Ontario.

OECD. (1996). *The Knowledge-based Economy*. Paris: OECD.

OECD. (2009). *Project Plan for the Adaptation of the CLA Instrument for International Implementation*. Paris: OECD.

OECD. (2013). *Assessment of Higher Education Learning Outcomes Feasibility Study Report: Volume 3: Further Insights*. Paris: OECD.

Office for Education Policy and Innovation. (2007). *Value-added Measures for School Improvement.* Victoria, Australia: Department of Education and Early Childhood Development.

Pellegrino, J., Chudowsky, N., & Glaser, R. (eds.). (2001). *Knowing What Students Know: The Science and Design of Educational Assessment.* Washington, DC: The National Academy Press.

Rooney, D. (2003). *Public Policy in Knowledge-based Economies: Foundations and frameworks.* UK: Edward Elgar Publishing.

Tal, B., & Enenajor, E. (2013). *Degrees of Success: The Payoff to Higher Education in Canada.* Toronto: CIBC Economics.

Tremblay, K., Lalancette, D., & Roseveare, D. (2012). *AHELO: Feasibility Study Report: Volume 1: Design and Implementation.* Paris: OECD.

World Bank. (2010). *Knowledge for Development: KEI and KI Indexes.* Retrieved from http://info.worldbank.org/etools/kam2/KAM_page5.asp