

Does Higher Education Teach Students to Think Critically?

Dirk Van Damme & Doris Zahner (eds.)

Paris: OECD Publishing, 30 August 2022. <https://doi.org/10.1787/cc9fa6aa-en>

<https://www.oecd-ilibrary.org/content/publication/cc9fa6aa-en>

Higher education contributes immensely to economic growth, social progress, and overall quality of life through the skills students and graduates acquire. Qualifications awarded by higher education institutions are valued because they are perceived to signal the skills required by labour markets and broader society. Employers use these qualifications as ways to identify and select job candidates who master essential and requisite skills. Higher education is trusted by employers and society to the extent that there is an equilibrium between skills supply and demand.

However, there are signs that the skills supply of graduates no longer matches skills demand in the labour market. Quantitative qualifications mismatch is turning into a severe issue in many countries, compromising productivity, growth and the continued increase in prosperity. Even more significant is the qualitative mismatch between the skills demand generated by the economic and social reality in labour markets and societies, and the supply of skills by higher education institutions. Employers and economic organisations express with increasingly louder voices that they are no longer confident that graduates have acquired the skills needed for the 21st-century workplace, in particular, generic skills such as problem solving, communication, creativity, and critical thinking.

Whether perceived or real, skills mismatch poses a serious risk to the trustworthiness of higher education. What is needed is more transparency about the skills students acquire. Unfortunately, this has not been a strength of most higher education systems. Transparency tools such as international rankings are quite good at capturing research-related measures or input measures in education quality but do not provide any insights into students' actual learning outcomes. The few available measures, for example, provided by the OECD Survey of Adult Skills (PIAAC), are far from sufficient and invigorate the demand for more and better metrics.

Between 2008 and 2013, the OECD led the Assessing Higher Education Learning Outcomes (AHELO) feasibility study. Despite a positive conclusion on the feasibility of the initiative, the proposal tabled by the OECD in 2015 to start the main study did not attract sufficient support and the project was abandoned. However, a small number of countries that supported the project decided to continue the endeavour at a smaller scale. The collaborative work concentrated on what was perceived to be the most interesting and urgent issue, i.e. the assessment of the generic skills of higher education students and graduates. The initiative found a partner in the Council for Aid to Education, Inc., a non-profit organisation in the United States with a long history of assessing generic skills in post-secondary education with its proprietary Collegiate Learning Assessment (CLA+) instrument. This volume reports on the work pursued between 2016 and 2021 to assess critical thinking and written communication, and associated skills in higher education institutions in six countries (the United States, the United Kingdom, Italy, Mexico, Finland, and Chile).

This report is a follow-up to the AHELO feasibility study and is one of the first international studies of generic skills proficiency in higher education institutions. It does not provide definitive answers but shows the power of assessing critical-thinking skills and how such assessments can feed into the policy agenda in higher education at national and international levels.

Its main conclusions can be summarized as follows.

Feasibility of critical-thinking skill assessment in higher education

Perhaps the first conclusion is the most important one: An international, comparative assessment of one of the most relevant learning outcomes of higher education is feasible. The chapters in the first part of this volume discuss in detail the construct validity, reliability and cross-cultural validity issues associated with an international assessment of higher education learning outcomes in the domain of generic skills. Their conclusion is clear: the CLA+ International instrument has potential as a valid and reliable assessment instrument. Of course, it is not the only assessment tool available on the market and it only covers a specific segment of relevant learning outcomes and skills but it has shown to function well in different contexts, across different systems and for various groups of students.

Moreover, the CLA+ assessment has empirically confirmed predictive validity on career and labour market outcomes later in life. Analysis of the US assessment data linked to surveys administered to employers and career advisors demonstrate that the critical thinking skills assessed by the CLA+ instrument are predictive of future educational success, career development and labour market outcomes. Assessments of relevant employability skills such as critical thinking provide more powerful indicators of human capital than measures of foundation skills such as literacy and numeracy.

That an international assessment is feasible was already the conclusion of the AHELO Feasibility Study, completed in 2013. But the CLA+ has further improved by learning from its implementation, analytical research on the data gathered, and the international collaboration of which this report is the result. Significant progress has been made on, among others, item development and, more recently, computer-based testing and computer-assisted scoring. This has greatly contributed to the useability and cost-efficiency of the assessment instrument, the processing of the data and the reporting of the analysis.

The experiences in the systems reported on in Part III of this volume and the methodological robustness of the instrument provide convincing proof that a wider implementation of the assessment in more institutions and systems is possible. A closer look into the substantive findings from the assessment in six systems in the next section will suggest that it is also worth doing.

Do students learn to think critically at university?

Part II of this report analyses the data of the assessment of over 120 000 students included in the aggregated database across institutions and systems. Of these, close to 100 000 were students in the United States. These students, almost equally split between those entering and exiting a first-degree programme, were assessed with equivalent versions of the CLA+ instrument over the period between 2015 and 2020. With the exception of Italy, all systems carried out multiple administrations of the assessment.

Across the sample, students entering a higher education programme on average performed at 'developing' mastery level of the test. Exiting students on average performed at 'proficient' mastery level. The shift is relatively small ($d = .10$) but significant. The total distribution remains more or less the

same across both subsamples, suggesting that the entire distribution moved to a higher score. The distribution is quite large, with on average 21% of students (average of country averages) performing at the lowest 'emerging' performance level. Thus, across countries, 20% of students performed at the lowest mastery level while 15% of students performed at 'accomplished' and 'advanced' mastery levels.

These general results across the six systems can be interpreted in different ways. Overall, it is encouraging to see that during their time in a higher education programme, students improved their critical thinking skills. However, given the importance that most higher education programmes attach to promoting critical thinking skills, the learning gain is smaller than could be expected. If universities really want to foster 21st-century skills such as critical thinking, they need to upscale their efforts. While universities produce graduates who can be considered, on average, as proficient in critical thinking, the distribution of achievement is quite wide, with one-fifth of students performing at the lowest level. With half of exiting students performing at the two lowest levels, it is difficult to claim that a university qualification reliably signals a level of critical thinking skills expected by the global market place.

The analysis cannot positively confirm that the learning gain is caused by the teaching and learning experience within university programmes. It is possible that, for example, selection effects (selective drop-out), general maturing of the student population or effects of learning outside university contribute to the average learning gain. However, the fact that the distribution in achievement remains more or less the same from entering to exiting students shows that the entire student population moves upwards, suggesting that the learning gain is caused by a common, shared learning experience.

Do background variables influence critical thinking skill development?

International large-scale assessments of learning outcomes such as PISA, PIAAC, Trends in International Mathematics and Science Study (TIMSS) and others show that variations in learning outcomes are quite heavily influenced by students' background such as gender, language, family background, migration status and other. It is interesting to examine whether this also is the case for the assessment of critical thinking learning outcomes at university.

The impact of language (whether students' primary language is different from the instruction/test language or not) shows to be statistically significant but rather small. Both entering and exiting students with a different mother tongue than the language of instruction perform slightly lower than other students except for the US sample of entering students. But, given the importance of linguistic proficiency for completing the test, these results are actually quite positive. Language barriers do not meaningfully hinder critical thinking.

This analytical finding from the entire database conceals contradictory findings in individual countries. With regard to language, it is interesting to note that in the United States sample, entering students with a different native language than the language of instruction performed better than students whose native language was the same as the language of instruction but that this advantage was reversed when they left the institution (see Chapter 6). In the England sample, both entering and exiting students whose native language was English outperformed students with a different native language (see Chapter 13). There are also interesting differences in the impact of language on the two different components of the test, the Performance Test (PT) and the Selected Response Questions (SRQ), which mobilise different language proficiency skills.

Gender does not seem to have a huge impact. There are some minor, statistically significant but, overall, small, differences between male and female students but no clear general pattern. However, as the country chapters in Part III illustrate, at a national level, gender can play a role. This is the case for Finland, where gendered patterns in scores by field of study can be identified (see Chapter 12).

The impact of family background has been examined through the parents' educational attainment variable. In contrast to language and gender, parents' educational attainment did show to have an impact on students' critical thinking performance, both among entering and exiting students and for both the international and the US sample. This points to a persistent effect of students' social-economic-cultural status on their educational achievement, even at this stage of their educational trajectory. Students from more disadvantaged backgrounds are more disadvantaged in critical thinking.

The question of whether parents' education also influences the learning gain between entering and exiting a university programme could not be answered in a conclusive way. The data suggest that students with higher parental educational status achieved a slightly higher learning gain than students with lower parental educational status. However, the relative impact of selection/attrition versus education remains unclear. More sophisticated research designs, preferably longitudinal, would be needed to answer that question.

Some of the country chapters in Part III explore other relevant background variables and their relationship with CLA+ assessment data. For example, in Mexico (Chapter 14) large differences were noted between students in campuses in metropolitan areas and students in campuses in remote, rural areas. In a country like Mexico, geography plays an important role through its association with economic development, social prosperity and levels of poverty and exclusion.

Do students demonstrate different levels of critical thinking by field of study and type of instruction?

Analyses reported in Chapter 8 show that there are significant differences in assessment of critical thinking skills between non-US students in programmes in different fields of study. On average across countries with relevant data, students in business and agriculture were found to have relatively low scores while students in the humanities, sciences and social sciences were found to have relatively high scores. This pattern holds for both entering and exiting students, suggesting a combination or interaction between selection and education effects. However, the highest learning gain achieved between entering and exiting university was found with students in health and welfare.

A more or less similar pattern was found for US students but with a slightly different ranking. Students in science and engineering were found to have the highest scores, both when entering and exiting, followed by social sciences and humanities.

Chapter 8 also reports on an interesting analysis of differences in CLA+ scores for different instructional formats for exiting non-US students. Although differences are relatively small, seminars, lectures and science laboratories are associated with the highest scores (with averages in the 'proficient' mastery level) whereas service learning and field work is associated with low scores. The positive result for lectures and negative result for service learning and field work contradict popular opinions on higher education pedagogy, which favour activating instructional formats. Critical thinking seems to flourish in instruction that requires deep engagement with content, as is the case for lectures, laboratories and seminars.

Differences between countries

Although the present study is not based on representative sampling within countries (except for Finland, which administered a system-wide assessment but with some institutions opting out), it is interesting to see whether there are meaningful differences in CLA+ results between countries. Chapter 9 analyses country-level differences in CLA+ scores for samples in five countries: the United States, the United Kingdom, Finland, Chile and Mexico. In the reporting of the results, countries were anonymised, as agreed with the country project managers in the study. Given limitations in the sampling, the impossibility of considering national data as reliable measures at the country level, and the agreement with participating institutions that they have full ownership of the assessment data, it was not possible nor desirable to rank countries. Still, some very interesting observations can be made. The comparison between countries is useful in exploring the importance of the national level as a relevant variable.

The data reported in Chapter 9 show clear variations in mastery levels between the five countries. Already, when students enter university, they exhibit very different proficiency levels in critical thinking. In countries A and D, half or more of the entering students scored at 'proficient', 'accomplished' and 'advanced' levels whereas in other countries half or more of entering students scored at the two lowest mastery levels. Large variation was also shown for exiting students. In country D, 70% of exiting students scored at the 'proficient' level or higher whereas in country C only 45% of exiting students scored at those levels. The average learning gain achieved while students were in university was unrelated to their mastery level when they started. Students in country C gained very little in proficiency even if they had started at a low level. Students in country D started at a much higher level but advanced a great deal before exiting while students in country E started at a much lower level but gained nearly as much. The results in country C are clearly disappointing. Other systems show greater levels of progress but country D makes clear that even with already high levels of critical thinking among entering students there is much opportunity to make significant progress during university.

The variability of CLA+ scores across country samples suggests that there are indeed significant differences between countries in the capacity of their education systems prior to and within higher education to develop critical thinking skills. As shown in the country-specific chapters in Part III of this volume, education systems differ in the policies, educational objectives and cultures within institutions. Still, in an increasingly global context for higher education institutions and, especially, for the employability and social participation of graduates, higher education systems that are better equipped to foster critical thinking skills will find themselves in a better place in the 21st-century environment. Given the changes in skill demand that are impacting all countries, though this varies depending on where countries sit in the global value chain, all countries should enable higher education institutions to perform better in fostering critical thinking.

That said, the variability in the country-level data is not high. International rankings have created the perception that quality differences between higher education systems are huge but the country-level data in this report contradict this. There are interesting country-level differences in the assessment of generic 21st-century skills but they are small and certainly do not mirror the steep hierarchical perception suggested by international rankings. In any case, with the exception of the United States and Finland, larger and more representative samples are needed for other countries before anything meaningful can be concluded about performance differences between countries.

Lessons learnt in participating countries

The experiences in individual countries illustrate that a shared interest in the importance of generic, 21st-century skills for employability and citizenship drives decisions to implement the CLA+ assessment. In all countries, strong political interest, often triggered by external stakeholders such as the business community, has been a necessary condition for moving ahead. When there is clear political consensus in favour of generic skills and a supportive political context for institutions, things start to move.

A second lesson learnt is the power of assessment to drive the reform agenda in higher education. The saying that only what is assessed matters, is true. Without assessment, debates on the importance of generic skills risk becoming partisan and divisive. With a credible evidence base, even imperfect, the debate is fuelled with data and becomes realistic.

A third lesson is about the importance of an inclusive approach. The institutional context in higher education is extremely important, implying that no government – let alone an international organisation – can impose an assessment on institutions. In all countries, institutional consent has proven to be a critically important condition for success. The failure of the AHELO project to move to a Main Study was probably due to the fact that conventional decision making at governmental level, without duly organised processes of discussion and negotiation with institutions, is doomed to fail in a higher education environment. Several country reports in Part III of this volume also point to the importance of motivating staff to positively support the assessment.

The fourth lesson learnt is about students. In a higher education environment, it is nearly impossible to force students to sit a test if they don't see the added-value for themselves. Several chapters in this volume explore the topic of student motivation and engagement. Suboptimal student motivation not only negatively affects participation rates but the quality of the assessment results too. Students are willing to sit the test and do their best if they view it as a reliable tool for their own interests. From this perspective, it is interesting to see the development of the CLA+ assessment towards rewarding successful students with digital badges and credentials. Prospective employers can access these to get an idea of the person's generic skills. This is a very promising development.

Prospects

It would be premature to call the present volume and the experiences in the participating countries a sufficient basis for moving to a large-scale assessment of generic, 21st-century skills in higher education. However, this volume illustrates the power of assessment to drive the policy debate on the importance of generic skills such as critical thinking. An important opportunity is opening for governments and institutions to develop initiatives to assess critical thinking, using the CLA+ instrument or others. Such initiatives will be powerful collective learning opportunities from which the entire global higher education community can benefit.