Abstract – This paper provides results from the first three years of a 4-year longitudinal assessment project on the development of cognitive skills using the Collegiate Learning Assessment (CLA+). The CLA+ is an online test, designed to measure cognitive skills (CS) (critical thinking, problem solving and written communication, with sub-scores reporting scientific and quantitative reasoning, critical reading and evaluation, and critiquing an argument). Cognitive skills are fundamental elements of engineering programs and central to the practice of engineering. The Canadian Engineering Accreditation Board’s (CEAB) requires programs to assess their student’s graduate attributes and have a system in place to use the assessment for curriculum improvement. CLA+ assessment constructs relate to the CEAB attributes of problem analysis, investigation and communication.

The testing was initiated as part of a strategy to track development of cognitive skills and to inform course improvement efforts. The testing was embedded in a range of engineering courses. Student achievement on written communication outcomes appeared to fall over time, and critiquing an argument was identified as an area of weakness, specifically for the first year students. Strategies were subsequently implemented as part of a curriculum improvement initiative. With the goal of maximizing success for all students, we investigated differences in achievement based on sex, language, and parental education, as well as the amount of effort put into the test. There were significant differences in achievement, prompting concerns about motivation and relevance to the engineering discipline. Students without English as a first language demonstrated negative gains over time. Parental education level was the strongest predictor of performance, but the most significant factor impacting on students’ scores was self-reported effort.

Key words: Assessment: Standardized test: motivation: cognitive skills: outcomes; program improvement.

1. INTRODUCTION

The ability to solve problems, think critically and communicate effectively is not only fundamental element of an engineering program, they are skills students’ find most valuable [1], and are essential in forging highly employable graduates [2]. Policy makers are placing greater importance on quality indicators. Cognitive outcomes and they are becoming increasingly important for institutions in demonstrating the quality and value of higher education. Governing bodies suggest we directly measure the value that institutions deliver to their students’ general and discipline-specific skills” [3].

Suskie [4] argues that there are five primary ways in which students learning can be evaluated, through:

i. Embedded course assignments
ii. Capstone experiences
iii. Field experiences
iv. Portfolios
v. Standardized tests

Embedded course assignments (i) generally focus on course goals, these indicators of student learning are often quite specific to the course and content being taught. Capstone experiences (ii) facilitate synthesis of knowledge through a program; they are generally applied in a summative manner, occurring toward the end of a program. Field experiences (iii) provide students with the opportunity to practice skills specific to a field. In many cases these are an excellent means way to determine the competency of students performing in particular disciplines, but they tend to be resource intensive and less suitable for large cohort assessment. The use of portfolios (iv) derives rich demonstrations of student learning, but can be time consuming to implement, and are difficult to directly assess. Unless implemented in an integrated manner, the above methods for evaluating student learning can have limited usefulness for longitudinal comparison, or comparison across programs. This paper focuses on the implementation of a standardized test (v), specifically the applicability for evaluating student learning across diverse groups and utility for informing course improvement efforts.
1.1 Standardized test

The Collegiate Learning assessment (CLA+) is a large-scale assessment tool. It was developed by the Council for Aid to Education, originally part of the RAND Corporation. In 2014 it was touted as a “game-changer for accountability in Higher Education” [5], but along with its advocates are many opponents [6]. Some institutions proudly publish their CLA+ institutional report, but the results are not always popular. Testing is often conducted as an add-on, happening at the institutional level rather than embedded within a course.

The CLA+ has two sections, the performance task section is designed to measure critical thinking, problem solving and written communication, and the selected response questions are designed to measure scientific and quantitative reasoning, critical reading and evaluation, and critiquing an argument. CLA+ require students to use a variety of sample material to address a problem in a constructed response format. The questions are modeled on “real-world” scenarios, but not specifically directed at the engineering field. As such, many instructors question their usefulness because of they are disconnected from the disciplinary context [7]. The testing was embedded in a range of engineering courses with varying levels of engagement from course instructors.

1.3 Purpose

Setting aside the larger question of external validity and the use of results for demonstration of value-add at the institutional level, the purpose of this paper was investigate:

i. Is the CLA+ an equitable test for students to track skill development?

ii. What lessons can be learned from the CLA+ about student learning for the purpose of course improvement.

In terms of equity for student differences, gender and English language status will be investigated. Parental education is commonly recognized as a predictor of academic achievement[8] so will also be investigated.

2. METHOD

2.1 Sample

The population comprised students enrolled across the 10 departments of the Faculty of Engineering and Applied Science. Ethics approval was obtained through the institutional research ethics board. Recruitment for the longitudinal sample coincided with the test scheduling. This began in fall 2013, when the students were in first year. Students were again recruited in the fall and winter of 2014-15 in their second year, and once more in the winter of 2016 when the students were in their third year. In first year, a research associate recruited students during scheduled course studio sessions.

The sample, including breakdowns for gender, language and parent education are shown in Table 1. There were 776 students enrolled at the beginning of first year. This number was fairly consistent over the first three years of the engineering program with only minor fluctuations due to transfer and enrolment status. The 30% of the population were represented in the consenting first year sample, this dropped to approximately 23% in second year, and 15% in third year. The proportion of male to female students each year was approximately consistent with the engineering average of 75% male 25% female. The breakdowns for language were also representative of the engineering population. There were no comparative data for the level of parent education.

<table>
<thead>
<tr>
<th>Year</th>
<th>Gender</th>
<th>Language</th>
<th>Parent Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>ESL</td>
<td>High School</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>College</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>ESL</td>
<td>High School</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>College</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>ESL</td>
<td>High School</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>College</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>English</td>
<td>Degree</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Post graduate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>English</td>
<td>Degree</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Post graduate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>39</td>
</tr>
</tbody>
</table>
In those cases, pizza was provided to encourage the student to attend.

The CLA+ was proctored through an online interface. The test allows a maximum of 90 minutes for completion; 60 minutes for the open-text response to a realistic situation “performance task” (PT) and 30 minutes for the multiple-choice or short answer “selected response questions” (SRQ) section. The CLA+ scoring is primarily automated; the total score is an average of the scaled PT score and weighted SRQ scores. The performance task responses are scored by an automated system using a validated rubric. There is no rubric for the selected response questions, only a generated score. In first year there were X test sessions. For the majority, the CLA+ means per test section and year.

### Results

The mean scores for first, second and third year are presented in Figure 1. In first year, the lowest sub-score (comparatively) was for “critique an argument”. There were drops in score for second and third year, the most notable being for the written communication sub-scores.

![Figure 1. CLA+ means per test section and year.](image)

The results were investigated by demographic breakdowns to determine if there were some groups who were performing better than others. There were no significant differences between performance of males and females, but there were significant results for language and parental education.

#### 3.1 Language

Student’s language was grouped as either English (as first language) or Other than English language (English as second or additional language (ESL)). Investigation found significant differences (see Figure 2). When considering language and year group, they appeared to be predictors of CLA+ total score ($F(2,530)=11.59, p<.001$), with an $R^2$ of .042. These results however, may have been more to do with the level of effort that the students put into the test rather than the actual student ability. Part of the exit survey for the CLA+ was a question in which student’s reported how much effort they put into the test. On a scale from one to five, where one is “no effort” and five is “my best effort”. The differences in student effort for language group are presented in Figure 3.

The effort in the ESL group dropped significantly between from second to third year. Correlations between effort and score were similar irrespective of the language status of students. For example, the relationship between selected response effort and score for ESL students was $r(72)=.406 p<.001$, the relationship between the same variables for the English students was $r(465)=.323 p<.001$.

#### 3.2 Parent education

Test takers reported the highest level of parent education as part the CLA+ exit survey. The responses were grouped for analysis as “high school or lower”, “College”, “Degree”, or Post-graduate”. Results are displayed in Figure 4. Differences were investigated using multiple linear regression and found that parent education and year group were also significant predictors of CLA+ total score ($F(2,524)=9.20, p<.001$), with an $R^2$ of .034. Student’s CLA+ score is predicted to increase by 28.77 points for each step of the parent’s educational level, with each year of study accounting for a difference 13.5 points.

Effort again confounds these results; the effect of effort on score generally decreases the higher the level of parental education. For example, effort and year group account for 14% of the Performance task score variance for parent education reported as “high school or lower”; 16% for “college”; 8% for “degree” and only 4% of the performance score variance is accounted for when parent education is reported as “post-graduate”.

Possible interaction between language and parental education was investigated with Pearson’s correlation; there was no relationship ($r=.062$ NS) between the groupings.
Figure 2. CLA+ Engineering 1st to 3rd year means by language

Figure 3. CLA+ Engineering 1st to 3rd year self reported effort by language

Note: The third year scores for “high school or lower” are based on a small sample so should be interpreted with caution.

Figure 4. CLA+ Engineering 1st to 3rd year means by parents education and time on task

4. Discussion

While recognizing that there are many ways to evaluate student learning, this paper focuses on the implementation of the CLA+ for evaluating student learning across the engineering student population, focusing on differences in achievement. There are challenges involved in any assessment of student learning, in the case of using the CLA+ it is important to consider that the questions do not specifically involve engineering problems, the test score did not count toward the course grade, and test was sampling student ability on one occasion each year (rather than in an on-going manner). Those conditions alone may have accounted for the lack of increase in mean scores between first and third year. These results were consistent with findings presented in 9, but it was puzzling that there appeared to be a decline in achievement over time. In first year, all the students who tested had their attendance rewarded with a course percentage mark. In second and third
year there were less incentives provided for students. Although there was no way to directly relate effort to the provision of an attendance mark for the test session, as fewer instructors offered an attendance percentage, the effort levels reported by students dropped.

It was reassuring to see that there were no significant differences between males and females, but there appeared to be differences based on language. There were minimal differences derived from language in first year, but performance (and associated effort) of ESL students dropped significantly in second and third year. In an engineering program, all students feel some level of pressure, but this must increase for students who do not have English as their first language. These results suggest that we need to undertake further investigation of the challenges ESL learners face.

Parent education statistics regularly correlate with other socio-economic (SEI) factors that influence student’s achievement [8], for example, a supportive family, conducive study environment and financial security. It is not surprising that students whose parents have higher levels of education did better in the CLA+ outcomes over time. It simply highlights the need to investigate how higher education can better support those who may not have as many natural environments advantages as others. It is almost a rhetorical question to ask if students of lower SEI backgrounds need to put more effort in just to improve their academic outcomes.

Understanding the limitations of student effort, the CLA+ data were still useful for informing course improvement efforts. This data, along with complementary assessment results, were used in changing the first year engineering course in which testing was done. New activities and scaffolding were developed to support critical thinking and problem solving development. Data collected over several years showed improved learning after these changes were made [9]. The results of these studies are also informing overall program improvement initiatives.

4.1 Limitations

If the CLA+ test results were included as part of the course grading the student effort would have been far more consistent. Reports for CLA+ data arrive months after each close of test window; even if the instructor was interested in (low-stakes) inclusion of CLA+ score in their grading scheme, the test-provider reporting delays prevent that from happening. Although the samples analyzed here were representative of the larger population, it would have introduced less error to analyze a dependent sample between year one and three. Due to consent and logistical matters it was not possible to collect are large enough dependent sample for analysis.

5. Conclusions

This paper provides results for the Engineering cohort from the first three years of a 4-year longitudinal assessment project on investigating the development of cognitive skills using the Collegiate Learning Assessment (CLA+). The constructs measured in the CLA+ align with those of engineering programs. The CLA+ was implemented to assess critical thinking, problem solving and written communication in a way that allowed longitudinal comparisons, to investigate student skill development over time and provide insight for curriculum improvement. Student effort declined over time, challenging the measurement of improvements in student learning. It appeared that written communication outcomes fell over time, especially for students without English as a first language. Parental education was a significant predictor for achievement, and the higher the parent education level, the less effect effort had on score. Effort levels withstanding there were results from the CLA+ that were used to inform course improvement.

ACKNOWLEDGEMENTS

The authors would like to acknowledge support from Higher Education Quality Council of Ontario (HEQCO).

References


