Lessons Learned from Using Competency Based Assessment (CBA) in a First Year Engineering Statics Course

Sean Maw, Shaobo Huang, Duncan Cree, Glyn Kennell, and Wendy James

University of Saskatchewan

Corresponding Author: sean.maw@usask.ca

Abstract – During the Fall term of 2020, the first year Statics course in the College of Engineering at the University of Saskatchewan was taught remotely, and synchronously, using a competency-based assessment (CBA) implementation. CBA is commonly used in other professions, especially medical education and teacher education, but it has yet to see widespread use in engineering education.

CBA can involve a number of key differences from conventional assessment practices. In this pilot, it involved the following key differences. Whereas previous versions of the course had involved assignments, labs, a midterm, and a final exam, each worth a certain weight in the overall course grade, the 2020 CBA version was broken into three modules, each with a “module test”. The module tests were superficially similar to a midterm, and there was no cumulative final exam. Open book assignments, quizzes, labs, and module tests consisted of questions and exercises that addressed a variety of learning outcomes (LOs) within the modules. The LOs were assigned weights in the overall course grade, as opposed to assigning weights to assessments themselves as in a conventional assessment system. Students could therefore overcome poor performances in early assessments of LOs, as better later results on the same LOs would replace the earlier results. A key feature of this approach was that students had at least two and typically three or more opportunities to exhibit competence with respect to the course’s LOs.

Another key aspect of this CBA implementation was the division of course material into three levels or “types”. Type A materials were the most basic building blocks of the course i.e. basic definitions, calculations, and concepts. Type B materials were basic integrative problems e.g. solving a basic 2D or 3D particle or rigid body equilibrium question, or solving a basic truss. Type C materials were advanced or “tricky” integrative problems that probed deeper understanding and required more adaptive problem solving. Students were required to meet competency thresholds for Type A and B materials i.e. they needed to exhibit a minimal level of competency in the LOs in order to pass the course.

Overall, the class excelled in this assessment format and anecdotal evidence suggests that the students enjoyed it. A summary description of the complete system will be presented in this paper, including how grades were determined, how assessments were conducted and

evaluated, how LOs were determined, and how the three levels of material were arranged. As well, basic statistical results from the class’ performance will be presented, along with a number of observations made by the instructors and some anecdotal impressions conveyed to them by students. The observed outcomes will be compared with the CBA literature for related STEM contexts, although the remote learning/COVID context did obscure the causes and origins of some of the observed outcomes. Changes that will be made in next year’s implementation of the course will also be discussed.

Keywords: Competency Based Assessment, CBA, Statics, First Year Engineering

1. INTRODUCTION

Assessment can play two fundamental roles in teaching and learning: summative assessment of learning that measures results when learning is done, and formative assessment to assist learning that helps check for problems and respond to them [10]. While some assessments can fulfill both roles, specific assessments typically focus on one role more than the other and furthermore, it is recommended that they do so [10].

In typical Canadian undergraduate engineering courses, it is common to see assignments, quizzes, midterms, and final exams where students solve problems or complete exercises to earn marks. These are classic summative assessments with grade weightings applied to each assessment e.g. “the final exam is worth 50% of the course grade”. Variations on this model exist where weightings can shift adaptively depending on performances in midterms, for example, but the key point is that grade weightings are assigned to specific assessments. Competency-based assessments (CBAs) typically do not work like this, as they emphasize results by learning outcome.

Well-designed multi-year professional programs have essential skills, processes, and conceptual understandings that they need students to master. Later courses depend on students developing specific skills and understandings in prerequisite courses. However, traditional assessment methods do not clearly distinguish between skills or understandings that a student is proficient with, in comparison to those where they may have gaps. Take a

final exam, for instance. The final exam grade gives an overall picture of performance. However, for mid-range grades, it is not immediately clear what has been learned or how well it has been learned. In addition, many skills or processes are only assessed once in traditional models, and if a student does not understand something, the course (and the student) moves on. The hidden gaps in understanding contribute to the promotion of students who will have difficulties being successful in later courses because they have insufficient preparatory knowledge or skills [5]. The use of competency-based assessment (CBA) and related scaffolds such as “top-ups” are specifically designed to result in more competent students and in a clearer understanding of what students can and cannot do well after course completion.

Competency-based assessment requires a clear set of learning outcomes (including knowledge, concepts, skills, processes, and sometimes dispositions), ideally across more than one course, scaffolded so that a progression in student learning is clearly laid out and agreed upon. In higher education in Canada, CBAs are most common in professional programs such as medicine that require accreditation and that have program learning outcomes [9]. Once overall goals for a year or whole program are agreed upon, each course is closely examined and considered through the lens of what is essential for students to know and do, in order to be successful in later courses or in the profession. These competencies are written as learning outcomes, starting with a verb that describes the level of complexity of the learning, and then including the key content and the context the learning must be demonstrated in [6]. The complexity of the learning described in the initial verb combined with the context of the learning outcomes matters because it clarifies what students will need to be able to do to complete later tasks, and it also makes clear what methods of instruction and assessment may be possible. A clear competence threshold should be established for each learning outcome in the course that is relative to that complexity. The outcomes and thresholds for competence are the basis for backward planning [2, 12] which address what is good evidence of success relative to the outcomes that have been established.

In traditional assessment, assignments, labs, tests, and projects are weighted and averaged together to create a grade. In CBA, the relative weight of each *learning outcome* (or LO) is established, and multiple types of evidence are associated with each learning outcome. Because the overarching goal is to determine student competence, early assessments are replaced with later evidence that is more complete or current rather than averaging all evidence together.

While CBA has been popular in Canadian professional programs such as medicine [8] or teaching [4], there is little evidence in the literature of its use in undergraduate engineering. One reason for this, which is inherent to CBA, is clear. More frequent grading is typically required

with CBA because multiple assessment sources of evidence are used [7]. Furthermore, not every student will follow the same path through a course, which increases logistical complexity in course delivery [1]. A third potential drawback is that students can get overwhelmed with the volume of assessments, especially if they are having difficulties with course materials.

On the other hand, the potential advantages are equally clear. With CBA, a passing grade means something valid and specific. CBA can also be a much more thorough system of promotional decision making. In addition, because it will typically involve coordination between courses, it can also do a more systematic job of preparing students for the challenges to come in upper years. Not coincidentally, it can work very well with validating performance against accreditation criteria e.g. [3].

It was with these potential benefits in mind that a pilot implementation was carried out at the University of Saskatchewan in the Fall 2020 term in the first year Statics course (GE 124 Mechanics I). The key question was: could CBA be implemented in a large (400+), foundational, first year engineering course to garner CBA’s benefits without incurring major costs?

Through this pilot, it became quite evident that the devil is truly in the details when one implements CBA. As such, this paper begins with a thorough description of what was done and why it was done. Then the results are described. This was not a formal study, so many of the observations are anecdotal. However, they have been partially validated in instructor focus groups and in student course evaluations. Basic comparisons with the previous implementation of the course are provided. However, attribution of cause and effect is impossible given that courses took place during the Fall 2020 term which was during the COVID-19 pandemic. Students were studying remotely and all assessments were open-book. These factors, among others, are covered in the Discussion. While few conclusions can be drawn with respect to CBA specifically, one thing is abundantly clear. Late-term performances of students on assessments were notably improved from previous years.

Please note that the University of Saskatchewan’s Behavioural Research Ethics Board was consulted about this quality assurance/program assessment work and they determined that no ethics review was required.

2. IMPLEMENTED CBA MODEL

Given its nature, CBA is inherently more complex than traditional assessment systems. Implementation details appear to influence outcomes in varied and nuanced ways. As such, academic analyses and comparisons of CBA are impeded by the variety of implementations. The intent here is to describe all of the details of this implementation. This will provide a fair basis for later comparisons against other implementations.

The implementation of CBA detailed in this paper took place in GE 124 (Mechanics I – Statics). There were approximately 400 students enrolled on September 1, 2020. The course was taught remotely, and initially asynchronously. However, after the first week of classes, the class switched to synchronous instruction with one professor giving a lecture via WebEx and PowerPoint, while another professor answered questions on the Chat channel of WebEx. The course was broken into three modules: Particle Equilibrium, Introduction to Rigid Body Equilibrium, and Topics in Rigid Body Equilibrium (e.g. trusses). Each module had its own assessments, instructor, and LOs, as well as weightings towards the final course grade. The course was delivered by four instructors, one for each module, plus an additional one who managed labs.

2.1. Type A/B/C Materials

Different levels of competency, and learning tasks of different complexity, should be assessed differently and given different types of feedback in CBA [1]. The classification system used in this pilot involved Type A/B/C materials.

Type A materials are fundamental definitions, concepts, and procedures that need to be automatized. The material is covered in class, is seen repeatedly, and is formulaic where procedures are involved. For example, a student may be given two vectors and may then be asked to calculate their dot product (in two different ways).

Type B materials are i) simple integrated problems, characteristic of the course, and/or ii) more complex conceptual questions covered in class. The processes may still be formulaic but they typically involve more steps. They integrate multiple Type A concepts and skills, and involve no significant extensions of concepts or procedures covered in class. For example, Type B material might require a student to solve a 2D rigid body equilibrium problem involving forces and moments.

Type C materials are expected to be challenging. Often, these questions will not have been addressed directly in class, and they will require deeper insights into concepts and/or unusual steps in procedures and/or recognition of key information that is uncommon. These problems require students to make sense of the unfamiliar, and to develop inferences in problem solving strategies. For example, a student may be asked what the maximum weight is that may be supported by three weight-bearing cords that each have tension constraints.

The main motive for developing this classification system was the belief that different levels of competence should be required for different types of material. Competence in Type A material should be very high and competence in Type B material should be quite high. All of this would be at the expense of specific performance expectations in Type C materials. In the development of USask's "re-engineered" first year program, which is the backdrop for all of the work in this paper, it was clear that

faculty were prepared to give up some performance at the highest problem solving levels if that resulted in more reliable and more consistent performance at the more fundamental and basic levels. The varying performance standards in the Type A/B/C materials would help ensure this outcome.

In short, the specifics of these standards were i) students had to pass all of the Type A materials, ii) students had to achieve at least a weighted average of 70% on Type B materials, and iii) students did not have to meet any performance standard for Type C materials.

What was trying to be avoided was the common problem in prior incarnations of GE 124 where students on the final exam would a) not be able to execute "Type A" procedures and/or b) would not be able to complete "Type B" problems. Yet, due to the nature of standard assessment schemes, it was frustratingly common to see students pass who had exhibited these outcomes. They passed by doing well enough on (parts of) other problems, but the gaps in their knowledge and skills were evident. Nevertheless, they would move on to second year with these deficiencies. This system ended up serving no one well.

2.2. Assessment System

With the complexity levels of material characterized by the Type A/B/C nomenclature, the next level of decision-making focused on logistics of assessment.

A core concept in CBA is the ability to have multiple tries to demonstrate competence. Furthermore, later tries that are successful can replace earlier tries that are not.

The Type A assessments were implemented in Mobius™ (<https://www.digitaled.com/mobius>) quizzes. These were regular weekly sets of quizzes that students could complete online generally in under one hour. Each of the course's three modules featured 3-4 "units" of Type A quizzes. The units corresponded to course-level LOs. Within each unit were a set of quizzes, each corresponding to "sub-LOs" (or SLOs). The SLOs provided systematic coverage of an LO at the Type A level. Every question in each quiz was drawn from a pool of similar questions, such that if a student needed to repeat a quiz, they would draw upon different questions of a homologous type. Students were given unlimited opportunities to pass the quizzes. With any incorrect answer, helpful feedback was immediately provided. Deeper tutorial assistance was also available if needed. Passing the quizzes constituted getting a perfect mark on short quizzes (e.g. 3 out of 3) or performing quite well on longer ones (e.g. 8+ out of 10). As these quizzes were being produced on an almost just-in-time (JIT) basis by a team of curriculum developers, the units tended to be released on a weekly basis, although in future course iterations it is expected that all Type A quizzes for a module will be available on the first day of a module.

Passing all Type A quizzes was required to pass a module, and therefore to pass the course. The Type A

material was therefore effectively pass/fail and did not affect numeric grades.

The Type B assessments were “typical” looking assignments, lab exercises, and test questions (on module tests at the end of each module). However, they were not typical insofar as an assignment’s grade did not have the usual meaning that it would have in a traditional assessment system. For the assignments in this course, each problem addressed specific LOs. These LOs were noted on the questions. For a given assignment, not all solutions would be marked and students would not know ahead of time which questions would or would not be evaluated. Before the course began, every Type B question in every assessment was mapped out to ensure that every LO was addressed at least three times during a module. Before the course started, instructors did not need to know which specific questions would be on each assignment and test. However, they did need to know the difficulty level and LOs that would be evaluated by each question.

Of note, students completed the assignments in pairs where cooperation was welcomed, while module tests were completed individually and remotely under open-book conditions and following strict time limits similar to conventional tests (with some extra time provided for uploading/downloading). Solutions were submitted to Crowdmark™ (<https://crowdmark.com/>) for evaluation.

LOs within each module were also weighted. If, after a module test, a student had not achieved at least 70% in the weighted calculation of Type B LO assessments for that module, they were given the opportunity to try again with a “Top Up” module test about two weeks later. Note that a pass/fail requirement for *each* LO could have been set. However, a weighted average *across* all LOs was adopted in a given module. As such, it was possible that a student could just barely meet the 70% threshold while not having done that well on one or more of the LOs within the module. This was a compromise that will be reassessed through further data analyses. The danger in having a pass/fail requirement on every LO is that the overall system can become too reductionist and unwieldy.

In either case, if a student did well on a certain LO on a module test, then that mark replaced earlier attempts on assignments for that LO. If a student performed less well on a module test for a specific LO, their module test mark for that LO was averaged with earlier grades for that LO from labs and assignments. If a student was under 70% on the Type B materials after the module test, they had to complete the Top Up module test two weeks later. If a student was still below the 70% threshold on Type B materials for a module after the Top Up, they failed the module on that basis (and therefore the course).

The Type C material was handled exactly the same way as the Type B material except for two critical differences: i) there were no “retries” on Type C material (these grades were simply averaged, per LO), and ii) a student could not fail a module based on Type C evaluations.

Ultimately, the weightings of the LOs and the weightings of the Type B and C material within each LO resulted in the following: if a student passed all Type A quizzes in a module, earned exactly 70% on their weighted Type B assessments, and earned zero on their Type C assessments, they would get a grade of 50% for the module. To get a higher grade, a student would have to do better than 70% on the Type B material and achieve some level of performance on their Type C material.

A passing grade in the course therefore overtly demonstrated sufficient mastery of the concepts, content, and skills needed for the courses in later years and provided specific information about which learning outcomes a student was successful at. The Discussion will expand on the “costs” of these gains.

2.3. Logistics

It is worth briefly describing the workflow that the instructional team settled into with this course, as it was somewhat different from conventional assessment systems.

The first and most important point is that the whole course had to be planned out in detail, ahead of time. This cannot be done “on the fly” with CBA. If a conventional assessment system is like a set of guitar strings, which can result in noise or pleasant music depending on how the independent elements are combined, then CBA is like a spider’s web. The reverberations of any change anywhere in the system are felt throughout the system. Few elements in the course are truly independent of each other.

What this meant was that the material covered in each assessment (level and type) was relatively inflexible. However, the choice of which problems could be “plugged in” to each assessment was still quite flexible. As such, one could characterize the workload as very front-end loaded. Once the course had commenced, preparing assignments and tests was actually easier than usual. One simply identified a specific question that met the LO and complexity specifications for a given question on an assessment, and that was that. Assignments, module tests, and Top Up module tests were all easy to prepare, on this basis, during the term. One simply had to figure out all the details ahead of time. Interested readers can contact the authors for documents that were used to plan the modules.

Another key logistics point was that Type B materials were marked by Teaching Assistant (TAs), while Type C materials were marked by instructors. Typically, a weekly assignment consisted of four questions (3 B’s and 1 C). Likewise, most module tests consisted of four questions (3 B’s and 1 C). Top Up module tests only included Type B questions, although to accommodate sickness and other reasonable absences, some students did complete a Type C question in the Top Up module tests.

3. RESULTS

This section describes how things went during the pilot. No systematic study was undertaken, largely due to the confounding effects of the COVID-19 pandemic, remote teaching, and open-book evaluations.

However, impressions of the instructors, anecdotal comments from students, other comments from course evaluations, and grading data, all suggest some fairly clear conclusions in some respects.

3.1. Instructor Observations/Experiences

CBA required more intensive planning of assessments prior to the start of the course. Once in the course, the system was not very flexible for changing assessments as there were too many ripple effects on a variety of grade calculations. Most instructors initially saw this as a disadvantage. However, later on in the course it was seen as an asset i.e. the system was operating on “autopilot”, and all the instructors had to do was “plug and play” good problems into assessments. This was seen as an advantage, especially as it made Top Up module tests simple to put together. Students also liked this organization of the assessments, as they would always know what to expect for any given assessment. Anecdotally, they regarded this as one aspect of the system’s inherent “fairness”.

During the course, it was very important to get marking done quickly as the CBA system depended on timely feedback on assignments. Typically, this meant returning Type B/C assessments back to students within one week. All instructors and TAs found this challenging at times.

One of the biggest unknowns in this pilot was how many students would need to do Top Up module tests. As it happened, approximately 10-15% of students needed to do each Top Up module test (the least for Module 1 and the most for Module 2). In practice, only 1/2 to 2/3 of those required to write the test did so, as some students had functionally left the course before officially withdrawing. Of note, approximately 40% of those writing the Top Up module tests brought their Type B marks up to a pass.

While the course enrollment was initially 400 students, after the course was fully complete (including all Top Ups), approximately 350 had passed. This level of attrition was comparable to past years, but the reasons for it were not, given COVID-19 and remote teaching issues.

One significant difficulty that instructors experienced, which can be remedied in the future, was software integration. This was the instructional team’s first use of the Canvas™ (<https://www.instructure.com/canvas>) learning management system. It was also their first use of Mobius™, as well as Crowdmark™ for some. Gradebooks were not able to be integrated in time, which made for a lot of extra administrative work. This was mostly invisible to students, as grades could be calculated offline and then uploaded to Canvas™ for student access. But better

integration would result in a markedly reduced administrative workload in future course iterations.

Finally, the instructors had the impression that the paired group work for assignments had been a failure. Although one goal of pairing students was to reduce the grading workload, every week there were dysfunctional teams that needed remedial interventions. However, in course evaluations, there were many comments made in support of the paired group work for such reasons as getting to know other students, helping each other, learning from each other, and developing professional interpersonal skills.

3.2. Student Feedback

One of the major concerns going into the pilot was how CBA would cope with student absences due to sickness or other reasons. As it happened, CBA turned out to be very helpful in this regard. This was especially evident in the remote teaching during the COVID-19 pandemic. Basically, if a student had problems, had missed an assignment or test, or was otherwise temporarily encumbered, CBA worked well for them. In most cases, absences simply reduced the number of opportunities to demonstrate competence. Generally, however, students could still get evaluated on LOs again. Therefore on a certain level, missing an assignment was not a problem. Several students noted that this really relieved their stress.

On the other hand, CBA is complicated (as the reader is no doubt aware by now). It was evident that most students did not really understand it at first. Indeed, it was evident that some students never understood it. However, those that did understand it by the end of the course generally expressed support for it in course evaluations. They really appreciated the “multiple attempts” aspect of the system.

To assist students in making sense of the system, a “grade calculator” spreadsheet was also provided for each module. Students could plug in grades by LO, as they earned them, and could even forecast what they would need to achieve on later assessments, if they wanted. Some students did use the grade calculators, and appreciated them, but most students did not appear to have tried them.

3.3. Type A Assessment Results

There were no students who failed GE 124 on the basis of failing just the Type A assessments. It was never the intent or expectation that students would only fail Type A assessments. Reassuringly, every student who failed a module’s Type A assessments also failed to meet the required Type B performance threshold.

That said, not all students found it equally easy to pass all of their Type A quizzes. This gets to a key issue with this implementation of CBA which is the division of material within the Type A/B/C levels. There were a few questions (among 1000) that were tougher than expected (they were conceptual Type B’s). Indeed, there were up to five questions that were released with errors. Mobius™

keeps track of “success rates” on questions i.e. how many tries were required before success. For these types of assessments, one does not want questions that everyone always gets correct (too easy; not worth doing). Nor does one want questions that very few students get on the first try (likely too difficult; Type B/C). There were about 10 questions that needed addressing for being too difficult (or unclear/incorrect) while there were only two questions that every student got correct on the first try.

3.4. Type B Assessment Results

Figure 1 shows the mark distribution for final Type B grades over the three modules. The small spikes at zero represent a number of students who had “checked out” of the course but had not deregistered. Note that students had to obtain at least 70% to pass in their Type B materials. The CBA approach to Type B questions required students to be able to successfully apply all key concepts and skills in the assessment with sufficient mastery. As Figure 1 demonstrates, the assessment design allowed the course to ensure that students were at the required levels of skill across all the major outcomes. In CBA, the goal is more students demonstrating sufficient mastery, rather than a normal distribution with many student missing the mark.

Fig 1 - Type B Grade Distributions by Module (N = 380)

3.5. Type C Assessment Results

Figure 2 is the same as Figure 1, except for Type C materials. Clearly there is a shift down in the distribution which is likely a function of a) no retries, and b) generally more challenging questions. This was not universal, however. It became apparent that some “Type C” material was really Type B level (and vice versa). The distribution of marks on Type C questions is also indicative of content that did not have to be mastered, but could be used to distinguish students between competency and mastery.

3.6. Overall Assessment Results

Figure 3 shows the final course grades for Fall 2020, compared to those from Fall 2019. While the numbers tell a clear story, the reasons for the differences are less clear. The students in 2020 definitely had higher grades. However, the assessments by any measure were easier.

Also, all assessments were open book in 2020, which was not the case for tests and exams in 2019. As well, it would have been easier to cheat in 2020, given the remote assessments, although the instructors did feel that cheating was not rampant due to a number of preventative measures. Overall, the instructors felt that the quality of the final work in all of the modules was superior to past years by a wide margin. While this is only anecdotal evidence, three of the instructors had taught GE 124 before, multiple times, and they were all in agreement that the quality of the final work in each module was superior to final exam results of past years.

Fig 2 - Type C Grade Distributions by Module (N = 380)

Figure 3 – Overall Grade Distributions Comparison

4. DISCUSSION

While the outcomes from this pilot appear promising, it is unfortunately impossible to say what led to them. It is likely that open book evaluations facilitated higher grades. Also, overall, the difficulty of tests was likely easier in 2020 with a greater focus on Type B materials. However, it is quite possible that the CBA format also helped boost performance. A related factor will be evaluated later this year, and that is the issue of long-term retention. Will this year’s cohort retain knowledge and skills better than previous cohorts? There is a study underway to help answer this question.

Remote teaching likely had positive and negative effects on different students. The live online lectures included “chat support” which was always busy. This was a positive feature unrelated to CBA that may have also helped students. However, the stresses associated with the pandemic likely contributed to lower performance [11]. The instructors noted that they answered more email inquiries than in previous years with the same number of students, which showed more engagement on the part of the students, potentially as a result of remote studies.

While instructors were concerned about how the pilot would run during COVID-19, CBA may well have been a positive for students during this time, as the forgiving nature of the system seemed to reduce stresses at a time when there were so many other stressors.

In terms of the course materials, the nature of the Type A/B/C materials was an issue that is likely to evolve for this instructional team. There is no clear dividing line between these levels of difficulty. Upon reflection, the instructors felt that some problems may have been misclassified. For a given LO, there were a couple of cases where performance on Type C materials exceeded those on Type B. This would suggest some misalignment. Also, there was a declining volume of Type A materials as one went from Module 1 to Module 3, which makes sense given the more applied and less basic content of later modules. There was also an open question of whether the Type C material in Module 1 was as difficult as Type C material in Module 3 i.e. do the levels shift somewhat with experience and exposure?

One aspect of CBA that was not implemented very appropriately in the pilot was the idea of Type A materials being completed before Type B materials were evaluated. There were no rigid deadlines for Type A quiz completion, to help ensure that no one failed due to being late on completing them. As a result, up to 30 students per module waited until the very end to complete them. Perhaps not surprisingly, these were typically weaker students who were most in need of completing them earlier. This issue will be rectified next year by requiring students to complete Type A materials before Type B materials will be graded.

Late submissions was another thorny issue. CBA does not handle the concept of “late deliverables” well. The approach that was adopted in this pilot was to generally not accept/grade late materials. CBA made this more feasible without too much negative impact on students, as students generally could re-address LOs in later deliverables. However, unless one makes “punctuality” an LO, CBA does not make it easy to deal with this behaviour.

As noted in the Results, student understanding of CBA was probably not very extensive or deep. Investments in concise and clear explanations are being made now. If CBA is implemented across multiple courses next Fall, it is expected that students will invest more time and energy in understanding it.

The value of Top Up module tests was also an interesting issue. They appear to have value for relatively stronger students who encounter isolated challenges in earlier work. However, for weaker students, Top Ups did not appear to have helped a majority, possibly due to sticking with ineffective study habits. This is an issue that will be focused on in the coming year.

Labs were another area where improvements could be made. There were four labs distributed among the three modules in GE 124. For the new first year program, program learning outcomes (PLOs) will be implemented. It became readily apparent in this pilot that lab skills and the ability to properly set up problems (with Given, Find, Assumptions, etc.) are PLOs. The question is, how can these be incorporated into a CBA scheme? The answer may be to create a 4th module for PLOs. This module will have a course weighting similar to the other modules. However, it will run in parallel with the other modules, whereas the first three modules run in series. Lab skills and other PLOs will be assessed throughout the course, as well as in other courses e.g. Dynamics. Students will receive direct course credit for them by course, but at the end of each term results will be compiled across courses for each PLO to provide students with broad feedback on PLOs.

It was noted earlier that one of the potential risks of CBA is “assessment saturation” i.e. too many evaluations. While unclear, this did not seem to be a problem for most students as the number of evaluations were similar to previous years. However, for students completing Top Up module tests, it did seem like it may have been a problem. In USask’s new first year program, time will be set aside for students to work on Top Ups, if they need to.

Finally, it is worth commenting on accreditation issues. CBA lends itself very well to supporting accreditation processes. This was evident as some LOs already align quite well with CEAB accreditation criteria [3]. Extracting this information for accreditation will be easier and more valid than previous efforts.

5. CONCLUSIONS

Based on the experience gained, it is believed that CBA is feasible and likely has tangible benefits in a large-class first year engineering context in Canada. Challenges are present, but they appear to be manageable. As expected, staff workload is the primary challenge. However, the flexibility in the system can work well for accommodating “life challenges” that students may experience and evidence suggests that students appreciate the forgiving nature inherent in CBA. The system is also inherently fairer in the sense that it gives students more time to show that they can do the work required of them.

While it is clear that students performed better in 2020 in GE 124 (as compared to 2019), it is not clear why. The explanation is likely a combination of factors, the division of which cannot be determined at this time.

Acknowledgements

The authors would like to thank the Association of Professional Engineers and Geoscientists of Saskatchewan (APEGS) for financially supporting this work.

As well, we would like to recognize the work of Zoe Mao, Nicholas Betancourt, and Caelia Gardiner in helping develop the Type A materials for the GE 124 Statics course, and their help with some figures used in this paper.

We would also like to thank Agnes d'Entremont of UBC for her work in linking ours to the BC OER initiative that has involved the development of our Type A materials.

We also want to acknowledge that this work was performed on Treaty 6 territory, the traditional lands of the Cree, Saulteaux, Stony, Nakota, Dakota, and Lakota, and the homeland of the Métis.

References

- [1] B. Abner, O. Bartosh, C. Ungerleider, and R. Tiffin, *Productivity Implications of a Shift to Competency-Based Education: An environmental scan and review of the relevant literature*. Toronto: Higher Education Quality Council of Ontario, 2014.
- [2] J. Biggs, "Enhancing teaching through constructive alignment," *Higher Education*, 32(3), pg. 347-364, 1996.
- [3] Canadian Engineering Accreditation Board, "2016 Accreditation Criteria and Procedures: Revised Feb 2017," *Engineers Canada*, 125 pp, 2017.
- [4] S. Elam, *Performance Based Teacher Education. What is the State of the Art?* American Association of Colleges for Teacher Education, 1971. Retrieved from <https://files.eric.ed.gov/fulltext/ED058166.pdf>
- [5] E. Holmboe, J. Sherbino, D. Long, S. Swing, and J. Frank, "The role of assessment in competency-based medical education," *Medical Teacher*, 32(8), pg. 676-682, 2010.
- [6] K. Kraiger, J.K. Ford, and E. Salas, "Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation," *Journal of Applied Psychology*, 78(2), pg. 311, 1993.
- [7] J. Lockyer, C. Carraccio, M. Chan, D. Hart, S. Smee, C. Touchie, E. Holmboe, and J. Frank, "Core principles of assessment in competency-based medical education," *Medical Teacher*, 39(6), pg. 609-616, 2017.
- [8] S.J. Lurie, "History and practice of competency-based assessment," *Medical Education*, 46, pg. 49-57, 2012.
- [9] A. MacFarlane and S. Brumwell, *The Landscape of Learning Outcomes Assessment in Canada*. Toronto: Higher Education Quality Council of Ontario, 2016.
- [10] J.W. Pellegrino, L.V. DiBello, and S.P. Brophy, *Chapter 19 – The Science and Design of Assessment in Engineering Education*: In: Cambridge Handbook of Engineering Education Research [A. Johri & B.M. Olds, eds.], New York, NY: Cambridge University Press, 2014, 763 pp. {ISBN: 978-1-107-01410-7}
- [11] B. Stolikier and K. Lafreniere, "The influence of perceived stress, loneliness, and learning burnout on university students' educational experience," *College Student Journal*, 49(1), pg. 146-160, 2015.
- [12] G. Wiggins and J. McTighe, *Understanding by Design (2nd Ed.)*. Alexandria, VA: Association for Supervision and Curriculum Development, 2005.