122nd ASEE Annual Conference & Exposition
June 14 - 17, 2015
Seattle, WA
Making Value for Society
Paper ID #11242

Efficacy of Using Grade Point Average to Predict Students' Cognitive Ability

Dr. James Warnock, Mississippi State University

James Warnock is the Interim Associate Dean for Undergraduate Studies in the Bagley College of Engineering at Mississippi State University. His background is in biomedical engineering and he has been a big proponent of self-directed learning and active learning in his classes and was the first person to introduce problem-based learning in the department of agricultural and biological engineering at MSU. James is also the Adjunct Director for training and instruction in the professional services department at ABET. In this role, Warnock oversees the development, planning, production and implementation of the ABET Program Assessment Workshops, IDEAL and the assessment webinar series. He also directs activities related to the workshop facilitator training and professional development.

Efficacy of Using Grade Point Average to Predict Students' Cognitive Ability in Bioengineering

Abstract

In a typical engineering course, student knowledge is assessed by periodic examination, usually administered as a mid-term exam or final exam. While this provides the instructor with some indication of what students know, it doesn't provide students an opportunity to learn the things they don't know. For courses that serve as prerequisites, students can progress to the next "level" with only having to know 60-70% of the course content. In contrast, in the video gaming world, the player has to achieve a perfect "score" in order to advance to the next level. If they do not achieve a perfect score they get another chance and so progression is often achieved through repeated attempts, especially at the higher, more difficult levels. The gaming, iterative approach was applied to a junior level biomaterials course, where progression was based on cognitive ability.

The course was divided into three separate modules; at the end of each module students were asked to complete three tests. The first test for each module consisted of 15 multiple-choice questions. These questions related to the *understanding* cognitive domain as defined by bloom's taxonomy. Students had to make 100% to progress to the next test, and they were allowed to repeat the test until they made 100%. The second test for each module was comprised of short answer problems that required students to calculate answers. These questions were designed to test the students' ability to *apply* their knowledge. Students that scored >90% were permitted to take the third test. Again, if they made less than 90% the test could be repeated. The third test consisted of poorly defined questions, where students were required to analyze raw data, interpret their results, apply them to the problem and provide a justification. This assessed *analyzing* and *evaluating* cognitive abilities.

The structure of this course prompted the following research questions to be asked: (i) Does student GPA correlate with the number of attempts a students needs to achieve 100% on each test? (ii) Do students with a lower GPA (i.e.<3.0) have the ability to master higher cognitive levels?

Data collected over two semesters did not show any correlation between student GPA and number of attempts to get 100% on tests. This finding was consistent across all different cognitive levels. Student GPA was also not a good predictor of cognitive ability, as students with lower GPAs were equally able to master application of knowledge as those with higher GPAs. Very few students were able to master evaluation of data and several students with high GPAs failed to make 100% on this test.

In conclusion, GPA is not a good indicator of cognitive ability and even students with a low GPA have the potential to learn fundamental knowledge and apply their knowledge to solve structured problems. A high GPA does not indicate an ability to function at the analytical or evaluation cognitive level.

Introduction

In a typical engineering course, student knowledge is assessed by periodic examination, usually administered as a mid-term exam or final exam. While this provides the instructor with some holistic indication of how much course content students have learned, it doesn't provide specific information about areas for improvement, nor does it offer students an opportunity to learn the things they don't know. For courses that serve as prerequisites, students can progress to the next "level" with only having to know 60-70% of the course content. In contrast, in the video gaming world, the player has to achieve a perfect "score" in order to advance to the next level. If they do not achieve a perfect score they get another chance and so progression is often achieved through repeated attempts, especially at the higher, more difficult levels. The gaming, iterative approach was applied to a junior level biomaterials course, where progression was based on cognitive ability.

Bloom's taxonomy[1] is familiar to many educators as a way of classifying different domains of cognitive ability. The original taxonomy developed definitions for six cognitive domains; knowledge, comprehension, application, analysis, synthesis and evaluation. The taxonomy was revised in 2001 to reflect relevance to the 21st century and the nouns originally used to describe the cognitive domains were replaced with verbs [2]. The categories are ordered from simple to complex and assume a cumulative hierarchy so that mastery of the lower domains is implied in order to master the higher domains [3]. Thus, students must remember and understand factual knowledge before they can apply, analyze or evaluate knowledge.

Thinking	Bloom's Taxonomy	Revised Bloom's Taxonomy
Highest	Evaluation	Create
	Synthesis	Evaluate
	Analysis	Analyze
	Application	Apply
	Comprehension	Understand
Lowest	Knowledge	Remember

The student grade point average (GPA) is universally accepted as an indicator of academic success. At the end of each semester, students receive a grade for each course they complete, which is converted to a numerical value (A = 4, B = 3, C = 2, D = 1, F = 0) and multiplied by the number of credit hours to give the number of quality points. Quality points are divided by the number of credit hours to determine the GPA. Although this method allows courses with different credit hours to be weighted differently, it does not allow for the relative difficulty of courses (i.e. lower level courses are treated as equal to upper level courses) and does not distinguish between subject areas so math, physics and engineering topics are weighted equally with general education electives. Recent studies have shown that raw GPAs systematically distort student achievement across majors [4]. Another problem with course grades is the grading policy in any given course is dependent on the individual instructor. For example, some instructors may give credit for activities that are independent of student learning, such as attendance or class participation. Furthermore, some instructors will grade with a fixed standard while others grade on a curve [5].

The objective of this study was to determine the efficacy of using student grade point average as an indicator of cognitive ability, as defined by Bloom's taxonomy. Specifically, two research questions were posed: (i) Can students with a lower GPA (i.e. <3.0) develop cognitive ability as it relates to the course content and (ii) (i) Does student GPA correlate with the speed at which they develop their cognitive ability?

Methods

Course Design

The *Biophyscial Properties of Materials* course was designed so that students would have an opportunity to demonstrate their cognitive abilities in relation to course content. The course was required for all biological engineering students. The goals of this course were (*i*) to provide the student with a fundamental knowledge of the major properties of materials used in biological systems, and (*ii*) to train the student to use this knowledge base to solve "real-world" problems in biological engineering. Specific learning objectives were:

1. To develop the student's understanding of terms and principles associated with materials engineering

2. To develop the student's ability to understand and explore the relationship between material properties and biophysical behavior/performance

3. To introduce the student to common natural materials

4. To develop the students ability to use his or her knowledge of material properties to make sound engineering decisions

To accomplish the goals the course was divided into three modules: Properties of Materials, Biodegradation, and Biocompatibility.

Participants

The study consisted of two independent student cohorts. The first cohort consisted of 48 undergraduate students in the junior year of the biological engineering program. Participants included 73% males (n = 35) and 27% females (n = 13), of which 17% were African American (n = 8) and 2% were international (n = 1). The second cohort was comprised of 57 undergraduate students in the junior year of the biological engineering program. However, six students failed the course and were consequently excluded from the study so as to not skew the data. Of the second cohort, 65% were male (n = 33) and 35% were female (n = 18) while 19.6% were African American (n = 10), 2% were Hispanic/Latino (n = 1), 4% were Asian (n = 2) and 2% were multiracial (n = 1).

Procedure

At the end of each module the students had to complete three tests. All tests were open book/open note and were administered online. Students were required to complete each test individually. The first test for each module consisted of 15 multiple-choice questions. These questions related to the “remembering” cognitive domain as defined by bloom’s revised taxonomy [2]. The first cohort of students was allowed an unlimited number of attempts for this test but they had to make 100% before they could take the second test. The procedure was modified for the second cohort, who were limited to five attempts and still had to make 100%. The questions were randomly selected from a set of questions so that students would not necessarily get the same questions when the test was repeated. As the questions were selected from the question database at random, different students did not get the same set of questions. This did not appear to pose any problems, as the questions had comparable levels of difficulty and none of them appeared to be answered incorrectly more frequently than others. Randomization of questions also helped to limit the opportunity for students to work together on their tests, which was not allowed. Furthermore, at the end of the test, students were given their overall score but were not told which questions had been answered correctly. The second test for each module was comprised of 10 short answer problems, randomly selected from a question set, that required students to calculate their answer and were similar to problems worked during class. These questions were designed to test the students’ ability to apply their knowledge. As with the first test, the first cohort had unlimited attempts, whereas the second cohort was limited to five attempts. Students that scored >90% were permitted to take the third test. Students received an online grade immediately after the test was submitted; however, because of limitations with the system to accurately grade answers, the instructor manually graded all tests and returned a score to the students within 24 hours of them completing the test. This minimized any delay if they needed to repeat the test. The third test consisted of poorly defined, open-ended questions, where students were given experimental raw data, required to analyze the data, asked to interpret the results, apply them to the problem and provide a justification for their choices. This assessed analyzing and evaluating cognitive abilities.

Analysis

A primary interest in this study was to determine if a correlation existed between student cognitive ability and cumulative GPA. By requiring students to make 100% on each test before they could progress, it was assumed that all students had the potential to develop the relevant cognitive skill. Therefore, we correlated the number of attempts it took each student to achieve 100% with GPA to determine if GPA could indicate the rate at which students developed their cognitive skills. To this end, least squares linear regression was performed with the statistical package SPSS Statistics, V21.0.

Results and Discussion

The study employed two student cohorts. The first cohort was allowed unlimited attempts on tests, whereas the second cohort was limited to five attempts. This change was made because we observed that some students in the first cohort would submit incomplete tests if they contained questions they couldn’t answer and repeat until the test contained questions they had answered previously and received full credit. By imposing a limit of five attempts students had to attempt all questions, as they could not predict which questions would be asked in subsequent tests. The limit was set at five attempts based on the median and mode values from cohort 1, which were both five for test one and were both four for test 2.

In the first cohort, the mean incoming GPA was 3.32 ± 0.58 and the median was 3.445. The lowest GPA was 1.86. 27% of students had a GPA below 3.0 and 43% had a GPA of 3.5 or above. In the second student cohort, the mean GPA was 3.35 ± 0.55 with a median of 3.53. The lowest GPA was 2.11 with 22% of students having an incoming GPA less than 3.0 and 53% having a GPA of 3.5 or greater.

At the first cognitive level, where students were required to remember information and answer multiple-choice questions, we expected all students regardless of incoming GPA to achieve 100%. Our data, presented in figure 1, shows that there was no correlation between GPA and the number of attempts required to make 100% on the test. In cohort 1 (figure 1A) where students were given unlimited attempts, the R^2 value was 0.004, indicating a 0.4% probability that there was a correlation between grade and number of attempts. For cohort 2 (figure 1B), where students were limited to five attempts, the regression line showed a decreased gradient, indicating there could be some correlation between GPA and number of attempts. However, the R^2 value was 0.104 showing a 10.4% probability that there was a relationship between grade and number of attempts.

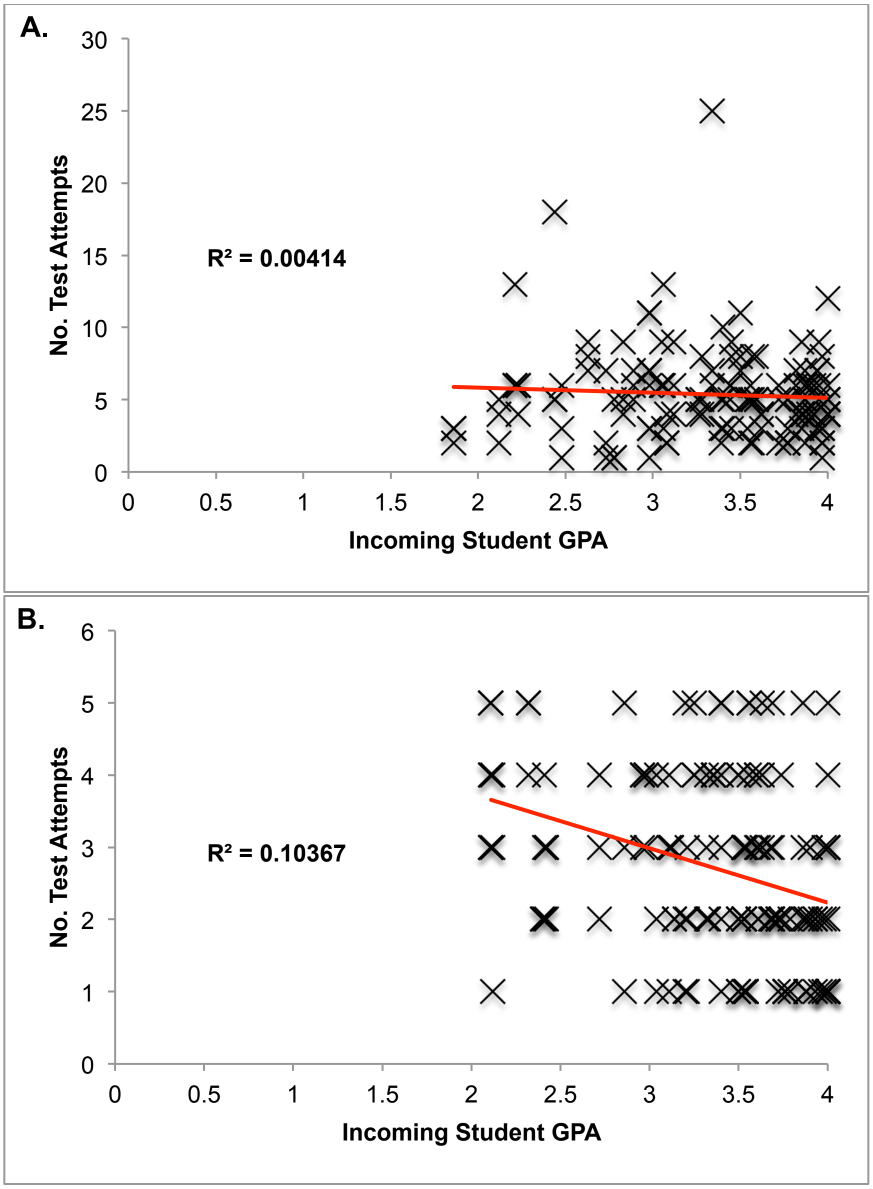

Figure 1. Number of attempts required on multiple-choice tests to make 100% vs. student grade point average at the start of the semester. Scatter plots show students from cohort 1 (N = 144) who had unlimited attempts (A), and students from cohort 2 (N = 153) who were limited to five attempts (B).

At the second cognitive level, students were required to solve short problem questions as a means of demonstrating their understanding and ability to apply their knowledge. From the first student cohort, eight students (16.7%) failed to make 90% or above on at least one of the three tests; eleven students (19.3%) from cohort 2 did not make 90% or above on at least one of the three tests they took after five attempts. The median GPA for students in cohort 1 was 2.91 and the mode was 2.50-2.99. In the second cohort, the median was 2.96 and the mode was in the range 2.50-2.99. Figure 2 shows the grade distribution of students that did not successfully complete test 2 as a percentage of the entire cohort. It was expected that there would be a more distinct relationship between the student GPA and number of attempts to make 90% on these tests as the presumption was made that students with a higher GPA had developed better cognitive skills for understanding and applying their knowledge. However, for cohort 1 who had unlimited attempts, no correlation existed and the regression analysis showed a 1.2% probability that there was some relation. For cohort 2, who were limited to five attempts, a week correlation was observed but regression analysis showed the probability of there being a relationship was only 2.2%.

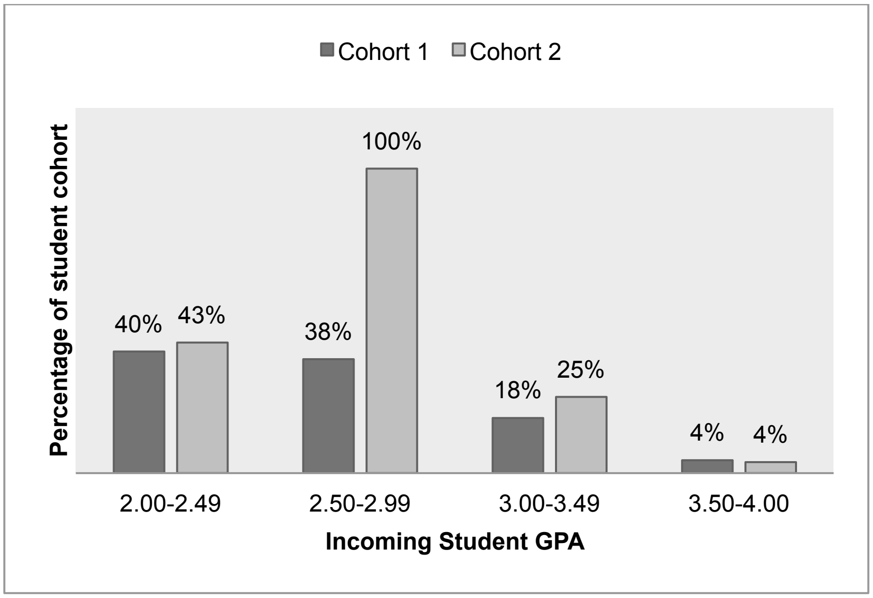

Figure 2: Percentage of students in each GPA range to **not** make >90% on the second test, assessing their understanding and analyzing ability.

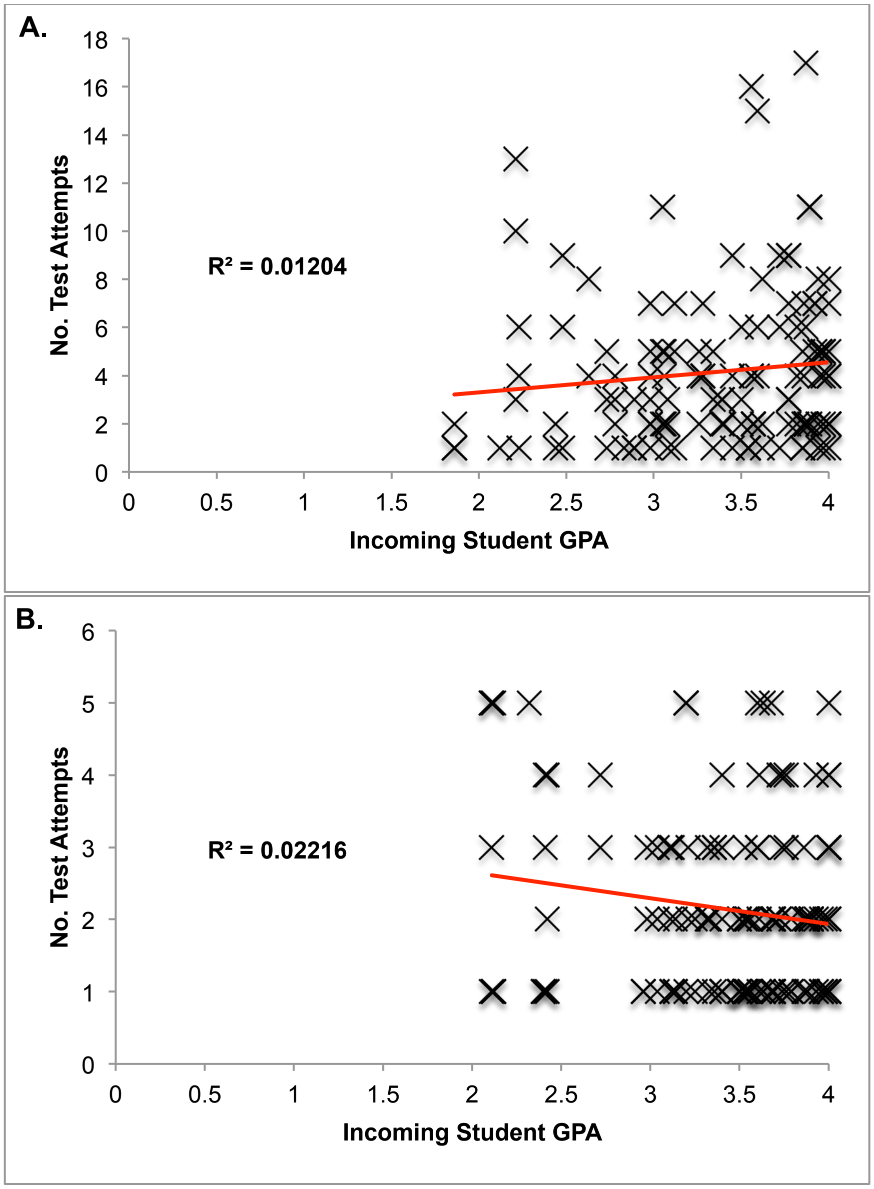

Figure 3. Number of attempts required on short answer problems to make 90% vs. student grade point average at the start of the semester. Scatter plots show students from cohort 1 (N = 144) who had unlimited attempts (A), and students from cohort 2 (N = 138) who were limited to five attempts (B).

For the final exam, students were given open-ended questions and supporting raw data as a means of determining their ability to perform at the higher cognitive levels, namely their ability to analyze and evaluate information. In cohort 1, 40 students took the final exam. Of those students, 21 (52.5%) were able to analyze and evaluate the data presented. For cohort 2, 31 students took the final exam. Of those students, 14 (45.2%) demonstrated the ability to analyze and evaluate data and use their knowledge to answer the questions sufficiently. Histograms showing the grade distribution of students successfully completing the exam are shown in figure 4A and histograms showing the grade distribution of students who did not complete the exam successfully are shown in figure 4B.

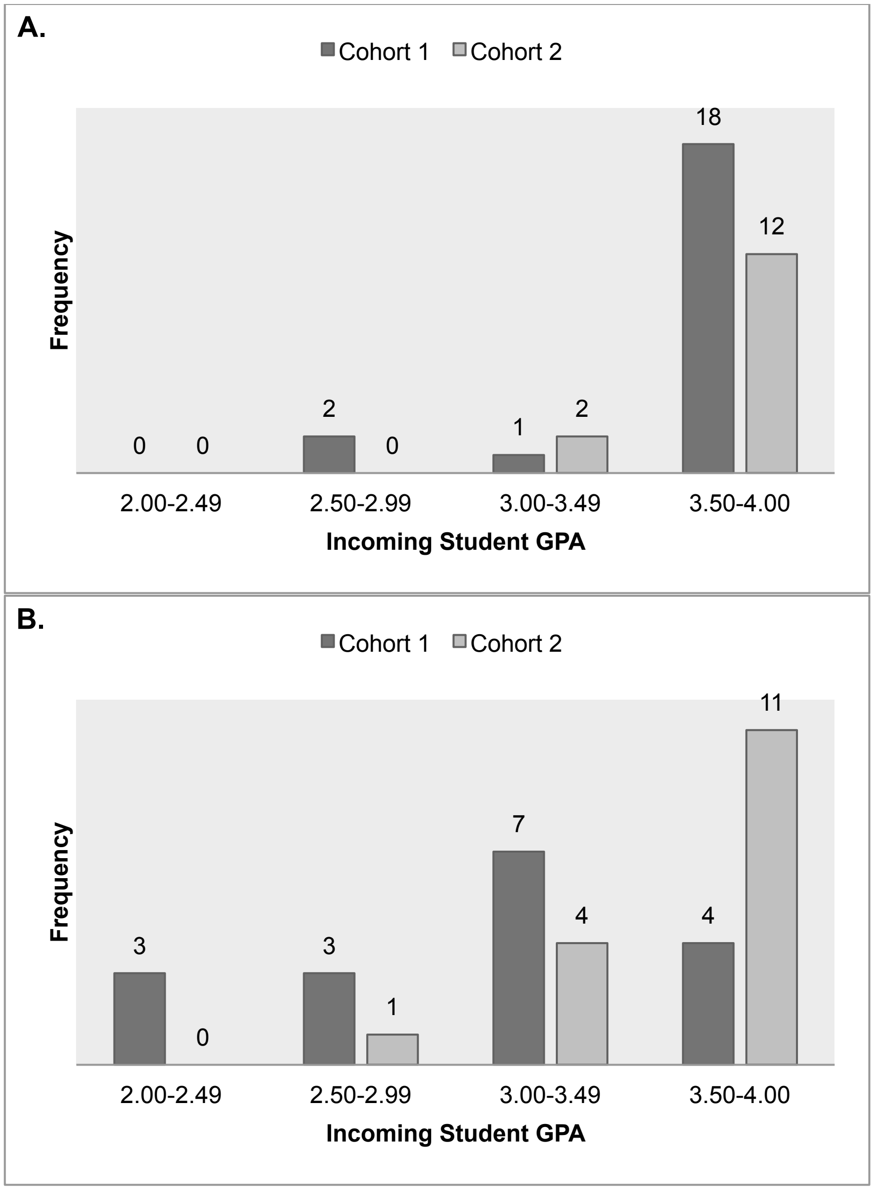

Figure 4: Histograms showing the grade range of students who successfully completed the final exam and demonstrated the ability to analyze and evaluate data (A) and the frequency of students who were unable to successfully analyze and evaluate data (B).

Summary and Conclusions

The grade point average is widely accepted as an indicator of academic success and readiness for professional engineering practice. However, our preliminary data suggest that GPA does not provide an accurate indication of students' cognitive ability.

The following observations were determined from our preliminary findings:

1. Students with low GPA have the same ability as students with a high GPA to develop their cognitive ability at the comprehension, understanding, and applying domains of

Bloom's taxonomy. There was no statistical correlation between student GPA and the number of attempts required to make 100% on the first two tests.

2. Students that demonstrate the ability to analyze and evaluate information have a high GPA (>3.50).
3. Although students that demonstrate an ability to analyze and evaluate information have a high GPA, not all students with a high GPA can analyze and evaluate information. Therefore, student GPA is not a good predictor of cognitive ability.

One issue that was not addressed in the study was student motivation. Students were not required to take the second or third test for each module, although these were linked to the final grade for the class. Some students may have self-selected to not take the final test, for example, because they were content to receive a B in the class. Other students may not have made 100% on the first or second test because of time management issues. Of the students that failed to pass any given module, the majority of them waited until a few days before the deadline to start and consequently ran out of time, not attempts. Of the students that did pass each module, they took their first attempt at least a week before the deadline.

The lack of correlation between GPA and cognitive ability indicates that univerity leaders should consider alternative ways of measuring academic success and using indicators that are a more acurate representation of students' abilities. One option could be to modify the GPA to a weighted GPA, where upper level course are given more weight than lower level courses. Alternastively, more weight could be given to engineering courses or major specific courses than general education and elective courses. A variation of this method is used by many medical schools who ask applicants their overall GPA and their biology, chemistry, physics and math GPA. An adjusted GPA could have a positive affect on engineering student retention as students would get more credit for taking harder classes and could reduce grade-induced student attrition[6].

Another option, albeit more contensious, would be to move award from a grade scale and allow students to focus on developing their knowledge and skills more than obtaining a letter grade in each class. This form of competancy based learning is being used in a few select institutions and could be a model for 21st century engineering education. For example, Alverno college, a liberal arts college in Milwaukee, WI, employs an ability based curriculum that allows students to develop their skills through ongoing assessment and feedback. Similarly, Harvy Mudd College in Claremeont, CA does not give first-year students grade; instead all courses are pass/fail. The rationale for this is slightly different, as it is intended to let students acclimate to college life without the stress of maintaining a GPA before they enter into the sophomore year where there is more intelectual rigor.

More research is necessary to determine how best to measure student academic success. This study merely indicates that what is currently accepted as an indicator for student ability may have limited utility and alternative options should be explored.

References

1 Bloom, B. S., College, C. o. & Examiners, U. *Taxonomy of educational objectives*. Vol. 1 (David McKay New York, 1956).

2 Anderson, L. W. *et al.* A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition. *White Plains, NY: Longman* (2001).

3 Krathwohl, D. R. A revision of Bloom's taxonomy: An overview. *Theory into practice* **41**, 212-218 (2002).

4 Bailey, M. A., Rosenthal, J. S. & Yoon, A. H. Grades and incentives: assessing competing grade point average measures and postgraduate outcomes. *Studies in Higher Education*, 1-15, doi:10.1080/03075079.2014.982528 (2014).

5 Rogers, G. Do Grades Make The Grade For Program Assessment? *Assessment 101: Assessment Tips* (2003).

6 Chang, M. J., Cerna, O., Han, J. & Saenz, V. The contradictory roles of institutional status in retaining underrepresented minorities in biomedical and behavioral science majors. *The Review of Higher Education* **31**, 433-464 (2008).