ASEE's 123rd Annual • Conference & Exposition • New Orleans, LA • June 26-29, 2016

New Orleans JAZZED about Engineering Education

Paper ID #14686

Academic "Predestination": Does It Exist?

Lt. Col. Phil Dacunto P.E., U.S. Military Academy

LTC Phil Dacunto is an Assistant Professor of Environmental Engineering at the United States Military Academy at West Point, NY. He earned a Ph.D. in the field of environmental engineering at Stanford University in 2013.

Academic “Predestination:” Does it exist?

Abstract

It often seems that instructors can predict who the best performers in a particular course will be by looking at their grades coming into the course. Those with the best grades coming into the course, the “good” students, usually seem to end up on top. However, does that relationship actually exist, or is it just a perception? Likewise, can we predict student performance in an engineering program based on their grades in certain classes (or the core curriculum as a whole) prior to entering the program? This paper seeks to answer those questions by analyzing grade data from several courses, and one engineering program. The course grade and program performance data came from 91 environmental engineering majors at an undergraduate teaching institution in classes graduating over a six-year period. Results of linear regression analysis of final course grades or program grade point averages (GPAs) against GPAs of the same students coming into the course or program are reported. In addition, the relationship between particular courses or sets of courses (for example, math and science courses only) taken previously and overall GPA in the major is explored. While all relationships were significant ($p < 0.005$), several were more useful in predicting future performance. A particularly strong relationship was found between an environmental chemistry course and overall performance in the environmental engineering major ($R^2 = 0.77$), due to both the course content and its lack of any group graded events; the relationships between overall performance in the major with 1st term GPA or 1st term math and chemistry grades were useful as well ($R^2 = 0.48$ and 0.52, respectively). Finally, 40% of students with 1st term GPAs less than 2.0 did not complete the full ABET curriculum, whereas all with 1st-term GPAs greater than 2.0 did. Understanding these relationships is important because by identifying students who may be at risk of performing poorly prior to or at the beginning of a particular course or engineering program, instructors and advisers can be ready to offer early assistance or objective evidence of how students with similar entering grades performed.

(1) Introduction

Students often seem almost “predestined” to receive a particular grade in a course or grade point average (GPA) in a major based on their past performance. The best students coming into the course or major often appear to be the ones that receive the highest grades, and vice versa. If true, this relationship between past and future performance in the classroom could be useful in advising engineering students.

Numerous studies have attempted to use previous performance to predict future academic achievement in engineering and other disciplines. For example, French, et al[1] employed variables such as high school rank and Scholastic Aptitude Test (SAT) scores to predict engineering students' GPA and enrollment status (i.e., whether they had persisted in studying engineering). They found several variables, to include high school rank and SAT math scores, to be significant in predicting both cumulative GPA and enrollment status. Likewise, Besterfield, et al[2] found that high school class rank and SAT math scores (in addition to other data reflecting attitudes toward engineering) were significant in regression models predicting persistence in the engineering discipline. Finally, Chemers et al[3] found that prior performance (high school GPA) was related to later academic performance, and that these past successes contributed to an attitude of "self-efficacy" (i.e., belief in one's own capabilities) that further contributed to success.

Several studies have used grades from specific courses, or sets of courses, to predict course grades or student achievement as a whole. For example, Huang and Fang[4,5] used eight variables including prior cumulative GPA, grades in pre-requisite courses, and grades on three mid-term exams to predict individual student final exam grades in an engineering dynamics. With a multiple linear regression model consisting of eight variables, they were able to predict the final exam grades to within 10% for 46-66% of the students. With more complex models such as a support vector machine learning model using 6-8 variables they were able to achieve accurate predictions for a maximum of 65% of the students. In addition, Ayan and Garcia[6] used up to six variables (first-year grades in math, biology, and chemistry courses as well as demographic factors) in both linear and logistic regression models to predict progress of students majoring in chemistry. With these somewhat complicated models, they achieved R^2 of 0.50 for the linear regression, and a correct prediction rate of 75% with the logistic regression.

The success of the aforementioned studies in predicting academic performance from past data suggests that the creation of such models is feasible, though they may need to be tailored to the specific institution or discipline. This study attempts to investigate such relationships for environmental engineers at our specific undergraduate institution; however, in contrast to some of the studies above, we attempt to find the simplest model possible, using linear regression to find single variables with the strongest relationship to overall performance in the major or in specific courses. These simpler models can be more helpful due to their ease of use; advisors can implement them simply by glancing at a student's grade data or transcript. For example, if the grade in one of the first courses taken in the major can be shown to be a reasonably good predictor of overall performance, advisors can easily use this to inform their student counseling.

At our undergraduate teaching institution, students take a broad core curriculum consisting of 26 semester-length courses from both the humanities and the sciences. Those choosing the ABET accredited environmental engineering major take an additional 17 courses in the engineering discipline to meet graduation requirements. With the Class of 2019, this changes to 24 in the core and 19 in the major. Traditionally, students chose their major in the first semester of their sophomore year (the third term), and began taking classes in the major in the fourth term.

As the result of an institution-wide curriculum redesign effort, beginning with the Class of 2019 students will now choose their major in the second term and begin taking major classes in the third term.

Thus, it would be useful to be able to determine if there are any metrics from the first term (e.g. academic performance overall, or performance in specific courses) that could be used to predict overall performance in the major, to include the likelihood of transfer to a less demanding (but non-accredited) version of the major which we offer for students who struggle with the ABET version. This "fallback" major requires three fewer courses but still allows the student to graduate. Such metrics would enable the early identification of at-risk students so that instructors and advisors can provide early assistance, or perhaps counsel them on selecting a different major.

Since the first two courses in the major (EV301 "Environmental Science for Scientists and Engineers," and XS391 "Environmental Chemistry") will now be taken in the sophomore year with the Class of 2019, it would also be useful to see if their performance in these courses can be used to predict future performance in the major. If so, advisors could present evidence to struggling students who would then still have time to adjust their priorities (for example, drop a club or sport that is taking excessive time from their studies), or change their major before it is too late.

For those students who are already in upper-level courses, the identification of useful metrics for predicting their grades in other courses can also be useful. Instructors and advisors can identify those students who may struggle before the class begins, and can be prepared to offer early assistance.

Thus, this paper hypothesizes that (1) overall academic performance in the first semester can predict overall performance in an engineering major; (2) performance in specific courses in the first semester, or early in the major, can predict overall performance in the major; and (3) performance in the engineering major thus far can predict grades in upper-level major courses. Understanding the extent of these relationships is important because by identifying students who may be at risk of performing poorly prior to the beginning of a particular course or engineering program, instructors and advisers can be ready to offer early assistance or objective evidence of how students with similar entering grades performed.

While the specific relationships reported are valid for this particular institution and engineering program only, they can be used in a general sense by advisors at other universities for the same purposes, especially with regard to the interpretations of causation vs. mere correlation. In addition, advisors can use these results to assist them in creating more specific relationships for their particular institutions.

(2) Methods

In order to explore the relationships identified above, grade data was obtained for 91 environmental engineering majors in classes graduating over a six-year period (2010-2015) at this undergraduate teaching institution. Classes graduating earlier than 2010 had a slightly different curriculum, and thus they were not included. Grade data were stripped of any personally identifiable information and included no demographics, but students were assigned an individual number so that their grades in specific courses or their GPAs could be compared. The data included both overall (including core classes) and major (only environmental engineering courses) cumulative GPA by term, as well as final grades in every course taken.

In addition, overall major GPAs through junior year from 22 majors of the Class of 2016 were obtained, as well as their final grades in EV481 "Water Resources Planning and Design," a senior design course. This data was used as an independent data set to verify one of the relationships established by the same data from earlier classes.

Data analysis and linear regression were performed in Microsoft Excel™, and final course grades were converted to a standard four-point scale. Table 1 explains in detail specific GPA and course relationships investigated, and to which hypotheses they relate.

Table 1. Relationships investigated in present study.

hypothesis	Independent variable	Independent variable notes	Dependent variable	Dependent variable notes
1	Term 1 GPA	Overall GPA	Final cumulative major GPA	n/a
2	CH101 grade	Final grade in 1st semester chemistry course	Final cumulative major GPA	n/a
2	MA103 grade	Final grade in 1st semester math modeling course	Final cumulative major GPA	n/a
2	CH101 & MA103 grades	Mean of final grades in both courses	Final cumulative major GPA	n/a
2	EN101 grade	Final Grade in 1st semester English composition course	Final cumulative major GPA	n/a
2	EV301 grade	Final grade in Environmental Science course (1st course taken in major, 27% group work)	Final cumulative major GPA	n/a
2	XS391 grade	Final grade in environmental chemistry course (2nd course taken in major, no group work)	Final cumulative major GPA	n/a
3	Cumulative major GPA through term 6	n/a	EV481 grade	Final grade in Water Resources Planning and Design, taken in term 7 (36% group work)
3	Cumulative major GPA through term 6	n/a	EV394 grade	Final grade in Hydrogeology and Hydraulic Systems, taken in term 7 (27% group work)

(3) Results/Discussion

Hypothesis 1: Overall academic performance in the first semester can predict overall performance in an engineering major

The relationship between the GPA in the first term and overall performance in the major (i.e., the cumulative major GPA upon graduation) is significant ($p < 0.005$), but not robust. The coefficient of determination (R^2) is 0.48, indicating that roughly half (48%) of the variability in overall performance in the major is accounted for by first term GPA. While a general trend is present (figure 1), there is significant scatter; a number of students who perform well (overall GPA >3.0) in the first term drop as much as a full point in major GPA by their graduation. Indeed, the slope (<1) and intercept (>0) of the regression line indicates that there is a general trend away from the extreme ends of the GPA spectrum, which is to be expected over the course of a student's college career. However, those students who perform very poorly in the first term

(GPA near 2.0) tend to remain toward the lower end of the major, and a number of students who do well in the first term continue to stay near the top.

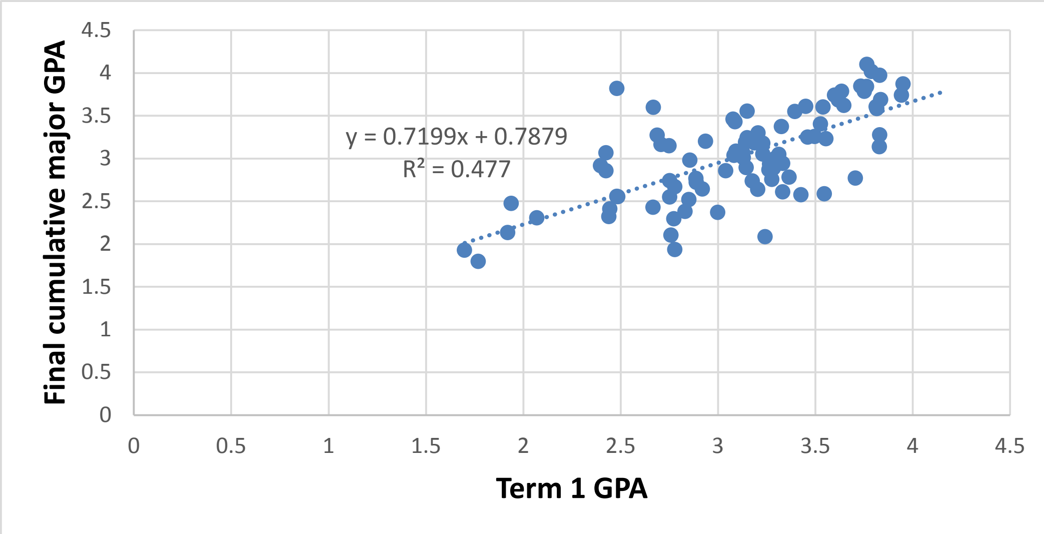

Figure 1. Relationship between overall GPA in the first term and final cumulative GPA in the environmental engineering major ($p < 0.005$, n=88; 3 of the 91 environmental engineers did not graduate, but the reason for separation is unknown). Note: no major courses are taken in the first term, and thus there are no issues with autocorrelation in this relationship.

This somewhat limited relationship between overall GPA in the first term and overall major performance is not surprising, since students are still adjusting to life and academics at the university in the first term, and classes taken include several in the humanities such as English composition and history (which are much different in content and requirements from engineering courses).

Only two of the 91 environmental engineers switched to the aforementioned non-accredited "fallback" major. Both had a GPA less than 2.0 after the first term, making this a possible "red flag" warning signal for use by advisors in counseling students in this situation. Notably, three others with GPAs less than 2.0 in the first term did not have to resort to the fallback major; however, none truly excelled, all graduating with major GPAs ranging from 1.9 to 2.5.

One significant shortcoming of this analysis is that it does not account for students who changed to an entirely different major from environmental engineering, since this data could not be obtained. However, these students are not included in the 91 total.

Hypothesis 2: Performance in specific courses in the first semester, or early in the major, can predict overall performance in the major.

Relationships between math and science courses taken in the first semester and overall performance in the major are significant ($p < 0.005$) but not particularly robust. For example, the R^2 value for the regression of the first semester chemistry (CH101) grade against overall major GPA was 0.43, and the relationship for a math modeling course (MA103) is about the same (R^2 =

0.45). Regressing each student's overall cumulative major GPA against the average of their MA103 and CH101 grades provided better results (R^2 = 0.52, figure 2). However, all of these were much better relationships than that between first semester English (EN101) and overall performance in the major (R^2 = 0.13, figure 3).

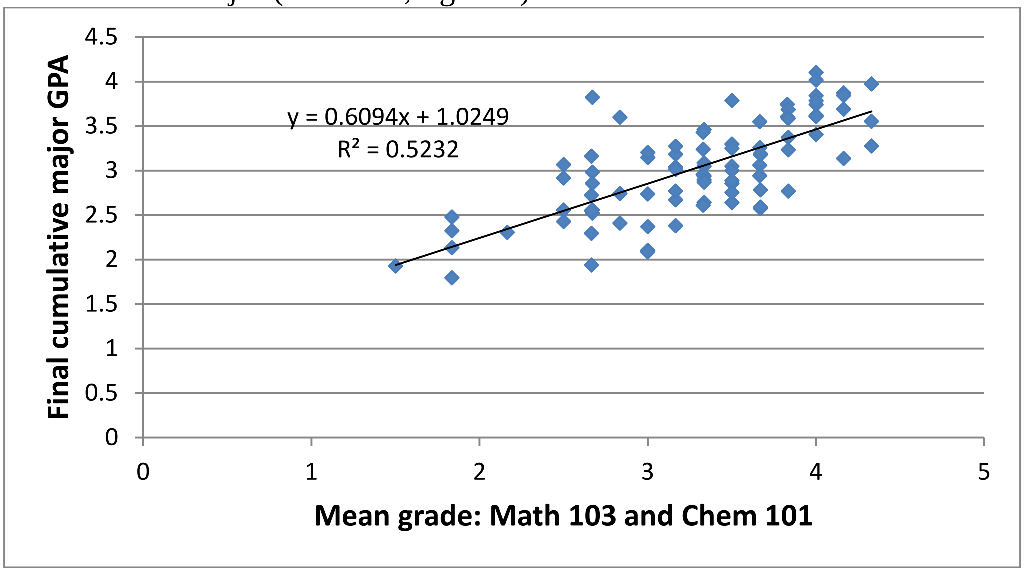

Figure 2. Relationship between final cumulative major GPA and the mean of each student's grades in freshman math and chemistry courses ($p < 0.005$, n=88). Course grades are converted to standard 4-point scale (A=4, B=3, C=2, D=1).

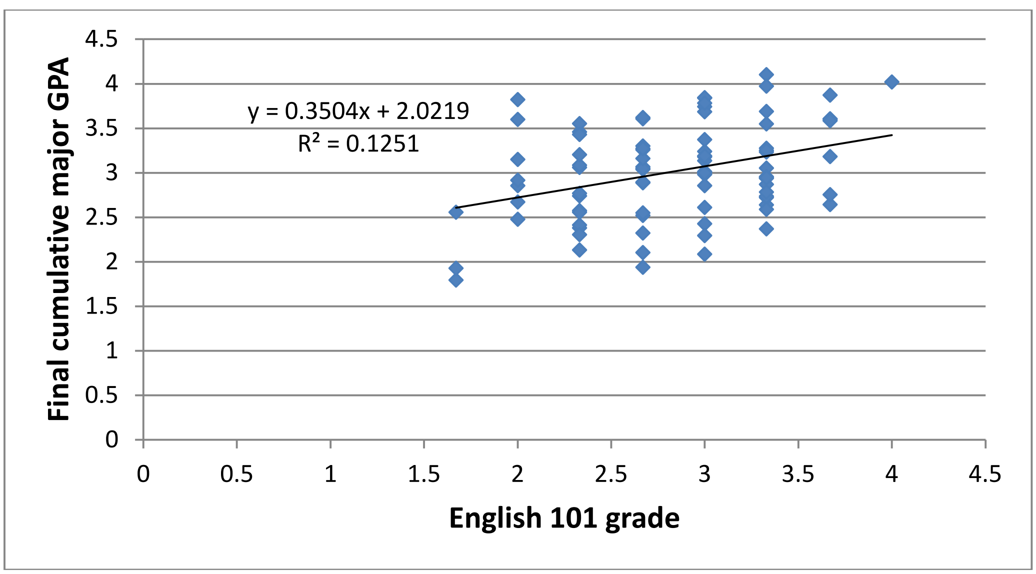

Figure 3. Relationship between final cumulative major GPA and freshman English composition grade ($p < 0.005$, n=84 since four students validated the course). Course grades are converted to standard 4-point scale (A=4, B=3, C=2, D=1).

This indicates that first-year math and science course performance may be a better indicator of future achievement in our engineering curriculum than humanities course performance. Since engineering relies heavily on math and science, this is not a surprising result in terms of causation, but the limited extent of the relationships suggests that we should not rely too heavily on the results from a single course, even if it is in the correct field.

Much better results were obtained when regressing final cumulative major GPAs against grades from the initial two courses in the major, EV301 Environmental Science and XS391 Environmental Chemistry (figures 4 and 5). Perhaps the best indicator is XS391 ($R^2 = 0.77$), since its content is fundamental to a number of future courses in the major. In addition, it does not contain the graded group work of EV301 (27%), which can muddle the results when trying to predict individual performance. Since XS391 will be taken in the fourth term for future classes, advisors can use this course as a potential indicator metric as well; in other words, students who struggle here may need to make some changes in their academic habits (or major) to ensure future success. At other institutions, advisors may be able to find a similar course which is taken early in the major, has limited group work, and whose content is fundamental to subsequent courses; such a course may be similarly useful in predicting performance.

Regression lines for all relationships investigated for hypothesis 2 had slopes < 1 and intercepts >0, indicating that there is a trend toward "the middle" as we compare performance in one course to performance in the major as a whole. In other words, students who earn particularly high or low grades in a single course tend to perform at a more moderate level overall in the major, which is to be expected as multiple course grades are averaged into the GPA.

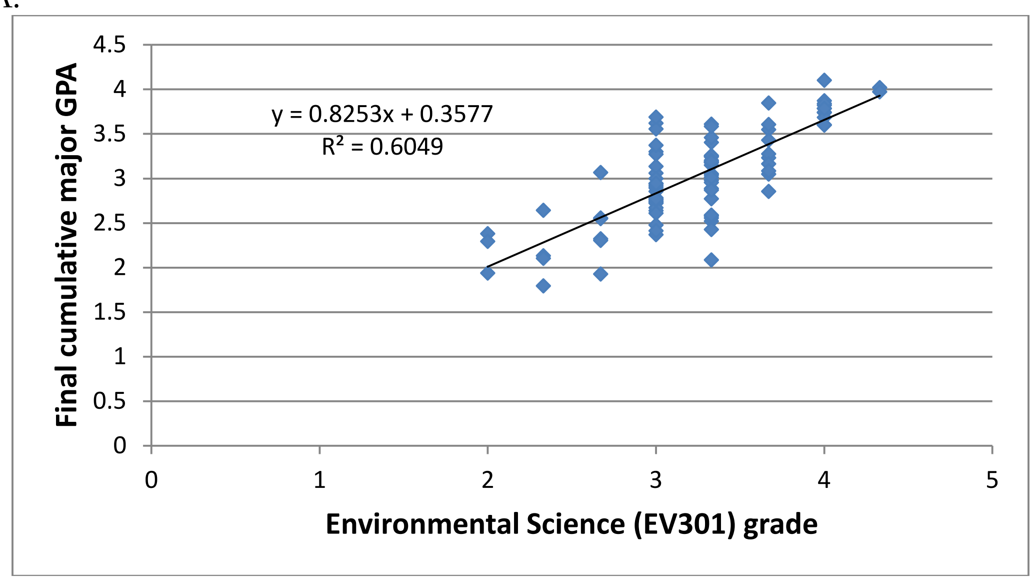

Figure 4. Relationship between final cumulative major GPA and Environmental Science grade ($p < 0.005$, n=88). Course grades are converted to standard 4-point scale (A=4, B=3, C=2, D=1).

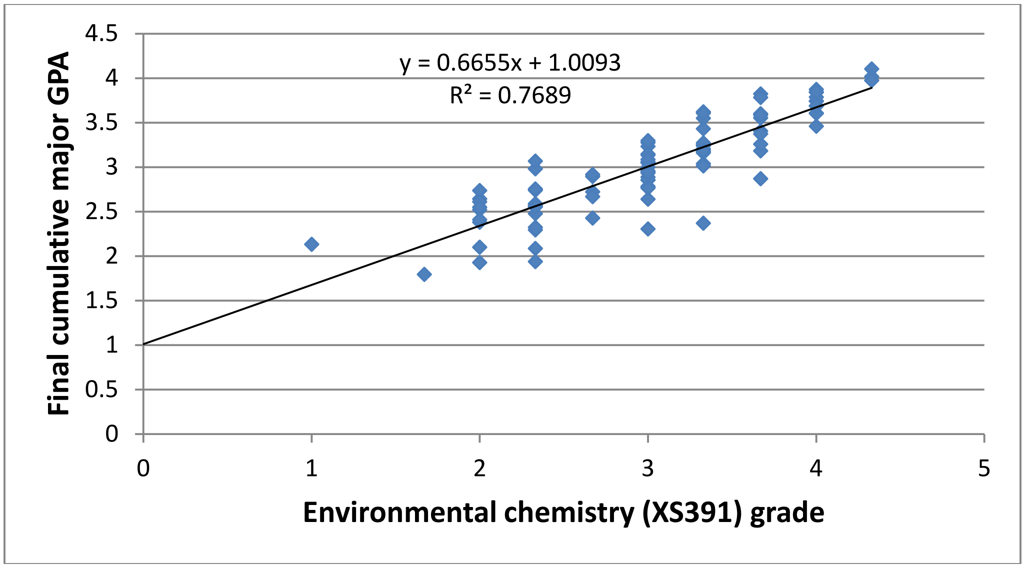

Figure 5. Relationship between final cumulative major GPA and Environmental Chemistry grade ($p < 0.005$, n=88). Course grades are converted to standard 4-point scale (A=4, B=3, C=2, D=1).

Hypothesis 3: Performance in the engineering major thus far can be used to predict grades in upper-level major courses.

The final area of investigation dealt with predicting performance in specific courses, versus the academic major as a whole. The focus was on two senior courses taken in the 7th term: Water Resources Planning and Design (EV481), and Hydrogeology and Hydraulic Systems (EV394). One important difference between the courses is the prevalence of group work: in EV481, it is 36% of the grade, whereas in EV394 it is only 27%. For both, the final grades in the course were regressed against cumulative GPA in the major after the 6th term (i.e., their cumulative major GPA coming into the course).

The relationship between cumulative major GPA in the 6th term and performance in EV481 is only moderately robust, with $p < 0.005$ and R^2 of 0.55 (figure 6). This relationship was used on an independent data set (environmental engineers in the Class of 2016) to predict performance in the course, and it served as a reasonably good predictor of final grades for the 22 environmental engineers enrolled (figure 7). The relationship was about the same ($R^2 = 0.60$, $p < 0.005$, n=88) with the hydrogeology course (EV394), indicating that the slightly smaller amount of group work in this course (9% less) had minimal impact.

Interestingly, both of these relationships also exhibited the aforementioned “trend toward the middle,” with slopes <1 and intercepts > 0. In other words, students with incoming GPAs at the upper and lower ends tend to perform a bit more toward the average. This may be due to the relatively high percentage of group work in these courses; students who would perform more poorly on their own may perform better when working with other members of a team, and vice versa.

These relationships indicate that while performance in the major overall is somewhat correlated to performance in future classes due to the overall nature of the work, individual courses within the major vary enough in terms of content and requirements that students can significantly influence their final grades.

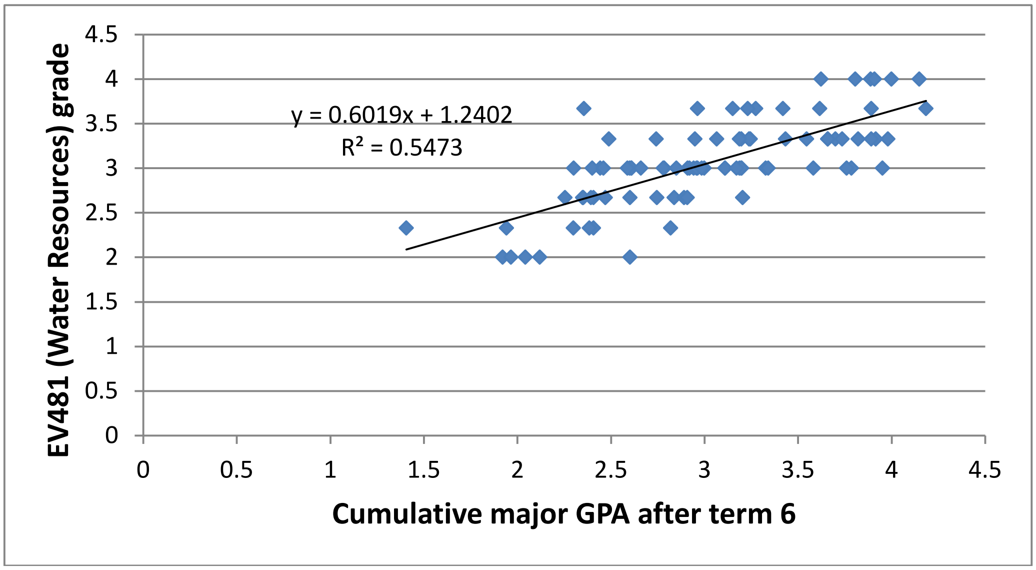

Figure 6. Relationship between EV481 (Water Resources) grade and cumulative major GPA coming into the course ($p < 0.005$, n=88). Course grades are converted to standard 4-point scale (A=4, B=3, C=2, D=1).

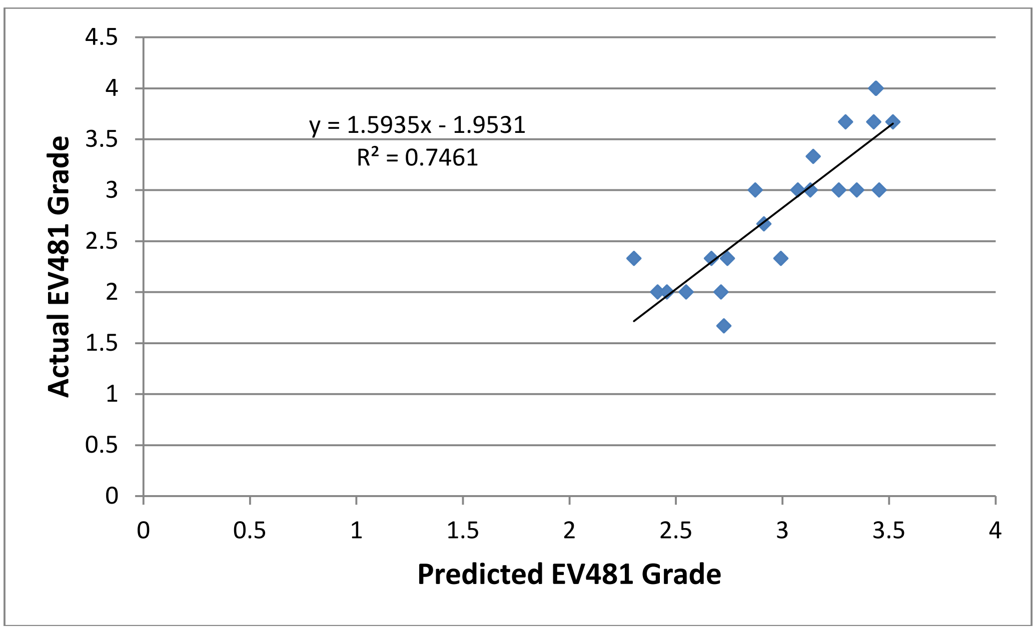

Figure 7. Actual vs. predicted Water Resources (EV481) grades for Class of 2016 (n=22), based on data from Classes of 2010-2015. Course grades are converted to standard 4-point scale (A=4, B=3, C=2, D=1).

(4) Conclusion

Analysis of grade data from 91 environmental engineering students from an undergraduate teaching institution indicates that overall academic performance in the first semester can indeed help indicate future overall performance in the major (hypothesis 1), though results are limited due to the wide variety of courses taken and the adjustment period inherent in the first semester of college. Thus, specific courses whose content is fundamental to the major and which contain minimal graded group work can be much more useful in indicating future performance (hypothesis 2). In our case, an environmental chemistry course taken early in the major met these requirements and was most useful in predicting performance in our program. Finally, cumulative performance in the major so far can help indicate performance in a course to a certain extent (hypothesis 3), though students certainly can perform beyond or short of their predicted grade due to the influence of group work and individual effort.

These simple models perform well enough that their advantage in terms of ease of use -- an advisor can implement them by merely glancing at a student's transcript -- outweighs any marginal gains from the more complicated multivariable models of studies mentioned previously. Indeed, the correlation coefficient (R^2) of the best single variable model (XS391, environmental chemistry) at 0.77 far exceeds those found by more complex linear regression models from Huang and Fang[4] (0.65) and Ayan and Garcia[6] (0.50).

Thus, it isn't just a perception: the stronger students coming into a course or major generally do perform better than the weaker ones. However, the requirements of specific courses vary, and students change over time, so we should be careful to not place too much confidence in our ability to precisely predict grades. Students are not necessarily "predestined" to success or failure; they can certainly influence their final grade within a letter grade or more up or down. However, it is true that students who are already struggling have a high probability of continuing to struggle, so the identification of potential "red flag" warning signals by advisors, such as a particularly low entering GPA or grade in a specific course, can be helpful in student counseling. Advisors at other institutions can apply some of these results to their own curricula, and possibly develop simple relationships of their own to assist students.

References

1. French B, Immekus J, Oakes W. An examination of indicators of engineering students' success and persistence. *Journal of Engineering Education.* 2005; 94(4); 419-425.

2. Besterfield-Sacre M, Atman C, Shuman L. Characteristics of freshman engineering students: models for determining student attrition in engineering. *Journal of Engineering Education.* 1997; 86; 139-149.

3. Chemers M, Hu L, Garcia B. Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology*. 2001; 93(1); 55-64.

4. Huang S, Fang N. Prediction of student academic performance in an engineering dynamics course: development and validation of multivariate regression models. *International Journal of Engineering Education.* 2010; 26(4); 1008-1017.

5. Huang S, Fang N. Predicting student academic performance in an engineering dynamics course: a comparison of four types of predictive mathematical models. *Computers and Education*. 2013; 61; 133-145.

6. Ayan M, Garcia M. Prediction of university students' academic achievement by linear and logistic models. *The Spanish Journal of Psychology*. 2008; 11(1); 275-288.