AC 2007-2750: CURRICULAR ASSESSMENT USING EXISTING ON-CAMPUS INFORMATION DATABASES

Andrew Kean, California Polytechnic State University

Andrew Kean is an Assistant Professor of Mechanical Engineering at California Polytechnic State University, San Luis Obispo. He received his Ph.D. from University of California, Berkeley in 2002 and his B.E. from The Cooper Union in 1997. His interests include energy conversion, climate change, air pollution, and sustainability.

Glen Thorncroft, California Polytechnic State University

Glen Thorncroft is an Associate Professor of Mechanical Engineering at California Polytechnic State University, San Luis Obispo. He received his Ph.D. from the University of Florida in 1997, with a research emphasis in Boiling Heat Transfer. His current activities focus on improvement of undergraduate laboratory education, including new experiments, instrumentation, and pedagogy in Fluid Mechanics and Thermal Sciences, as well as introducing Uncertainty Analysis into the undergraduate curriculum.

Curricular Assessment Using Existing On-Campus Information Databases

Abstract

Assessment of engineering program success is critical for continual improvement. While this assessment can take many forms, this work outlines an underutilized method of indirect assessment that takes advantage of already existing campus-wide information databases. Most university campuses have some form of information database which contain student records, course records, and/or faculty records. The methodology of using these databases to assess program performance is motivated by the popular book "Freakonomics" by Levitt and Dubner (William Morrow, 2005). While somewhat limited in depth, the scope of questions which can be answered with the databases are only limited by the creativity of the analyst. Of particular interest to the authors are the trends in student grades for key courses (e.g., statics and thermodynamics) over time, as department personnel have changed significantly. Also, we were curious to see connections between success in a prerequisite course versus a follow-up course. This work outlines some of the obvious and not so obvious assessments that are possible, as well as identifies potential pitfalls the analyst should avoid.

Introduction

It is the goal of most engineering education programs to accomplish continual improvement. To address this goal, assessment of the program's success in achieving stated learning outcomes is necessary. For this reason, ABET Criterion 3 for 2006-2007 Accreditation Cycle requires identification and assessment of program outcomes. Extensive efforts to improve assessment in education, and specifically engineering education, have already been performed (e.g., Astin, 1991; Shaeiwitz, 1996; Ewell, 1998; Pelligrino, Chudowsky, and Glaser, 2001; Olds, Moskal, and Miller, 2005). Generally, program assessment at the department level can be a time-consuming and expensive effort for the members of a department. Faster and easier methods for assessing program success and improvement would be welcome by many engineering departments.

This manuscript describes an underutilized method of assessment based on existing campus-wide information databases. The indirect assessment methodology is motivated by the popular book "Freakonomics" by Steven Levitt and Stephen Dubner (William Morrow, 2005). This text presents a novel way of discerning how people behave in the real world. The authors analyze existing datasets to obtain unexpected answers to some creative questions. Their analysis is based on the two key concepts: 1) that human behavior is strongly influenced by incentives, and 2) the conventional wisdom is often wrong.

With these fundamental concepts in mind, we present a methodology for the specific application of assessment of engineering programs. Ewell (1989 and 1998) has pointed out previously that capitalizing on existing data is a key approach for assessment implementation. The hope of the authors of the present work is to provide a useful technique for understanding the performance of our students and faculty better.

Most university campuses have some form of information database which contains student records, course records, and/or faculty records. While vast, these datasets are typically underutilized for the purposes of assessment. One key reason for this is that sufficient attention is not always paid to ensuring data collected at the institutional level are made available to individual academic units (Ewell, 1998). A second reason may be that curricular assessment with existing databases does not easily fit into existing categories of assessment (e.g., those discussed by Olds, Moskal, and Miller, 2005). This work outlines some of the obvious and not so obvious assessments that are possible, as well as identifies potential pitfalls the analyst should avoid. Some results for California Polytechnic State University, San Luis Obispo are presented.

Methodology and Results

When doing any such "data-mining" project, it is important to have prepared appropriate questions in advance. If you simply look at large datasets looking for any possible correlations, it is possible to find "statistically significant" results that are due simply to random fluctuations. Instead, it is better to approach the database with specific questions in mind. The keys for developing appropriate questions are found in the fundamental concepts of Levitt and Dubner. Specifically, we would like to draw your attention to the second concept: that conventional wisdom is often wrong. This far reaching concept has proven a most useful tool for developing questions to be investigated using on-campus databases.

California Polytechnic State University, San Luis Obispo has historically used IBM "SIS/Student Information Systems" environment for recording student records. In the very recent past, the campus has begun transitioning to a PeopleSoft "HRSA/Student Administration" environment. The methods presented here are obviously appropriate for any database environment, but the analysis results presented were obtained using PeopleSoft. All analyses were performed using manual data query techniques. While somewhat laborious, the authors' used manual queries for this pilot study to make the technique approachable to the maximum number of people. Automated queries would add an additional level of depth to the approach, but require greater knowledge and/or training to perform. A primary goal here is to determine whether advanced techniques are warranted.

Use of on-campus databases can provide assessment of success at many levels of detail; from California's overall higher education system on down. For instance, the methods could easily be used to compare academic preparation of transfer students from 2-year junior colleges compared to continuing students. These types of issues are really beyond the influence of a single academic department, so instead we have chosen to focus here on using databases to assess department-level issues.

Of particular interest to the authors are the trends in student grades for key courses (i.e., Statics and Thermodynamics) over time. In the last 10 years, our department has seen significant turnover in faculty. The general perception is that the new, younger faculty are easier graders compared to the senior faculty they replaced. Student evaluations are a major part of our University's tenure-track review process, and the argument is presented that younger faculty give easier grades in order to positively influence the reviews from their students. We were interested to see if this accusation of grade inflation is verified with data for the Mechanical Engineering

Department at Cal Poly. To address these questions, we used PeopleSoft to lookup grades in ME 211 Engineering Statics and ME 302 Thermodynamics from 1990 through 2006. Due to the sheer number of students in these "service" courses (i.e., ME courses taught to many different departments), we decided to focus specifically on Fall Quarter and retrieved records for years 1990, 1991, 1995, 1996, 2000, 2001, 2005, and 2006. This allowed us to identify variability over both short and long periods of time without requiring an overwhelming effort. The trend in average GPA given for Statics for these years is given in Figure 1.

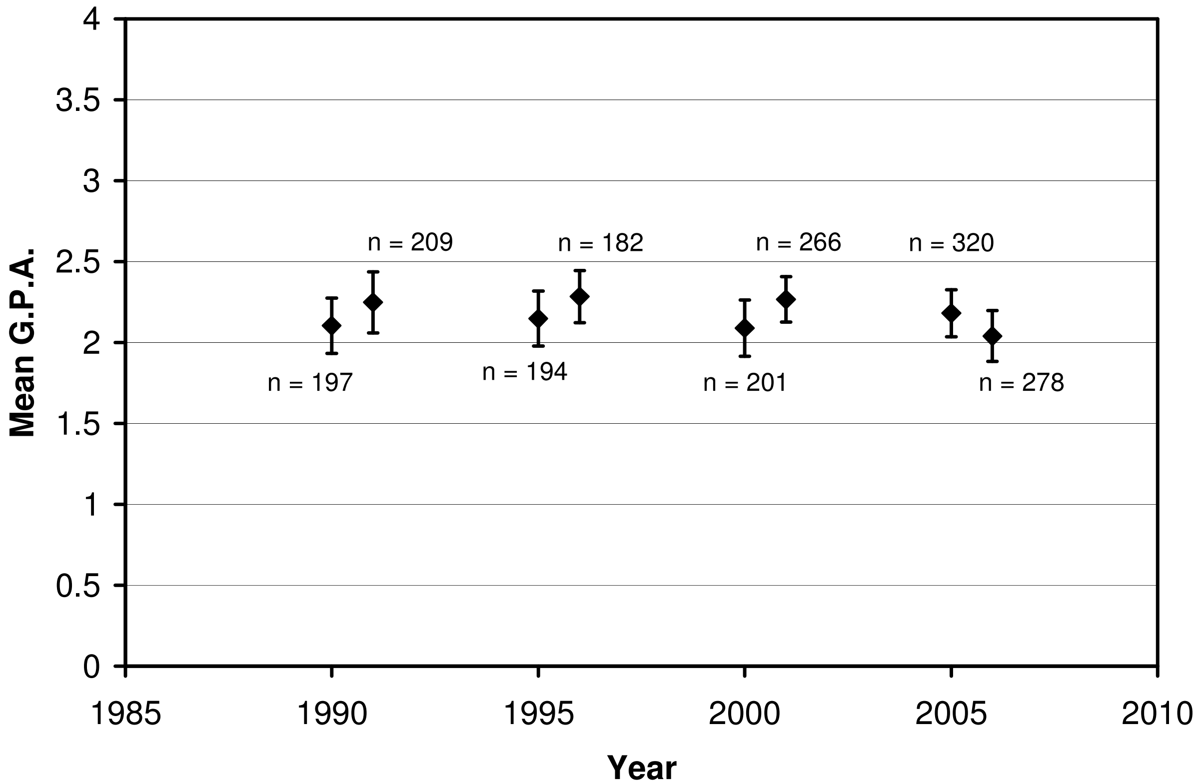

Figure 1. Trends in Engineering Statics Grade Point Averages (± 95% C.I.) Over Time

Note that the short term variability exceeds the long term variability with no significant trends over time. This indicates that the conventional wisdom regarding grade inflation by junior faculty is not apparent for this specific course. Our Statics course has what we refer to as a common final, whereby all faculty teaching the course collaborate on both preparation and grading of the final exam. It occurred to the authors that this common final could effectively act to control grade inflation for this course. On the other hand, for our Thermodynamics course, each faculty member is responsible for preparing and grading their own final. This lack of a common final could encourage grade inflation. Figure 2 presents average GPA given for Thermodynamics for the terms previously discussed.

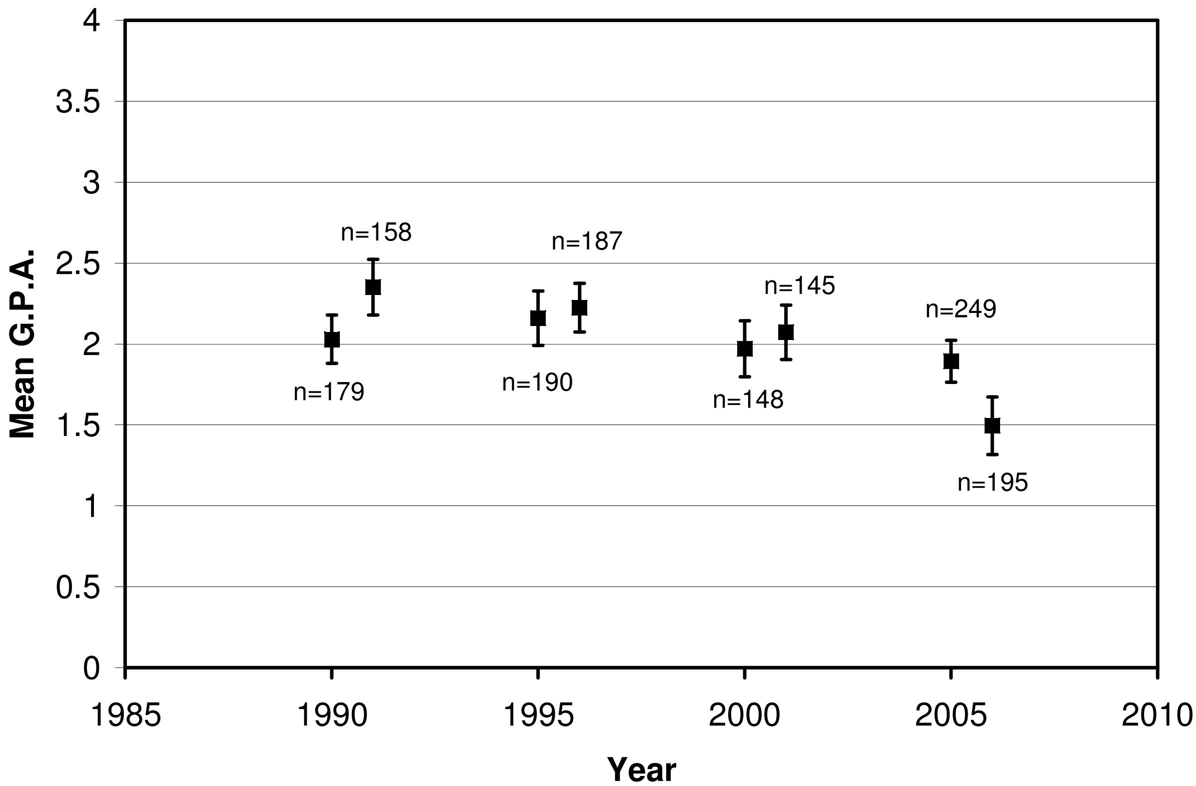

Figure 2. Trends in Thermodynamics Grade Point Averages (± 95% C.I.) Over Time

Comparing the two previous figures, it is clear that both the short term and long term variability in grades given in Thermodynamics exceeded those for Statics; the common final exam has the effect of providing greater normalization from year to year. It was quite surprising to us to find that rather than grade inflation, we may in fact have severe grade deflation occurring in Thermodynamics. Being rigorous educators is admirable, but it is unlikely that the student performance was dramatically different between Fall 2005 and Fall 2006. Having such drastic GPA swings from term to term is likely a disservice to our students and should be addressed by our department.

Upon completion of Thermodynamics with a passing grade (D- or better), ME students must also take the follow-up course (Thermal Engineering), though not necessarily in sequential academic terms. We wanted to use our campus database to determine if success in the prerequisite was related to success in the follow-up course. While success in a course is not easy to define, for our purposes here we initially defined success by what grade was earned. Whether grades are an adequate means to evaluate learning is discussed elsewhere (e.g., Rogers, 2000) and will not be specifically addressed here. Our analysis is based on three academic quarters in 2006 of Thermal Engineering grades (n=274, taught by 2 faculty members). By investigating the two years prior to 2006, we were able to capture the Thermodynamics grade of 90% of those 274 students (n=248, taught by 9 faculty members). Figure 3 compares grades in Thermal Engineering with those from Thermodynamics by student. Because course grade is a discrete variable with limited possible values, our ability to see trends in the Figure is reduced. Schools with no +/- grading would have even less resolution. We found that there is almost no correlation between grades received by individual students in these two courses ($R^2 = 0.08$).

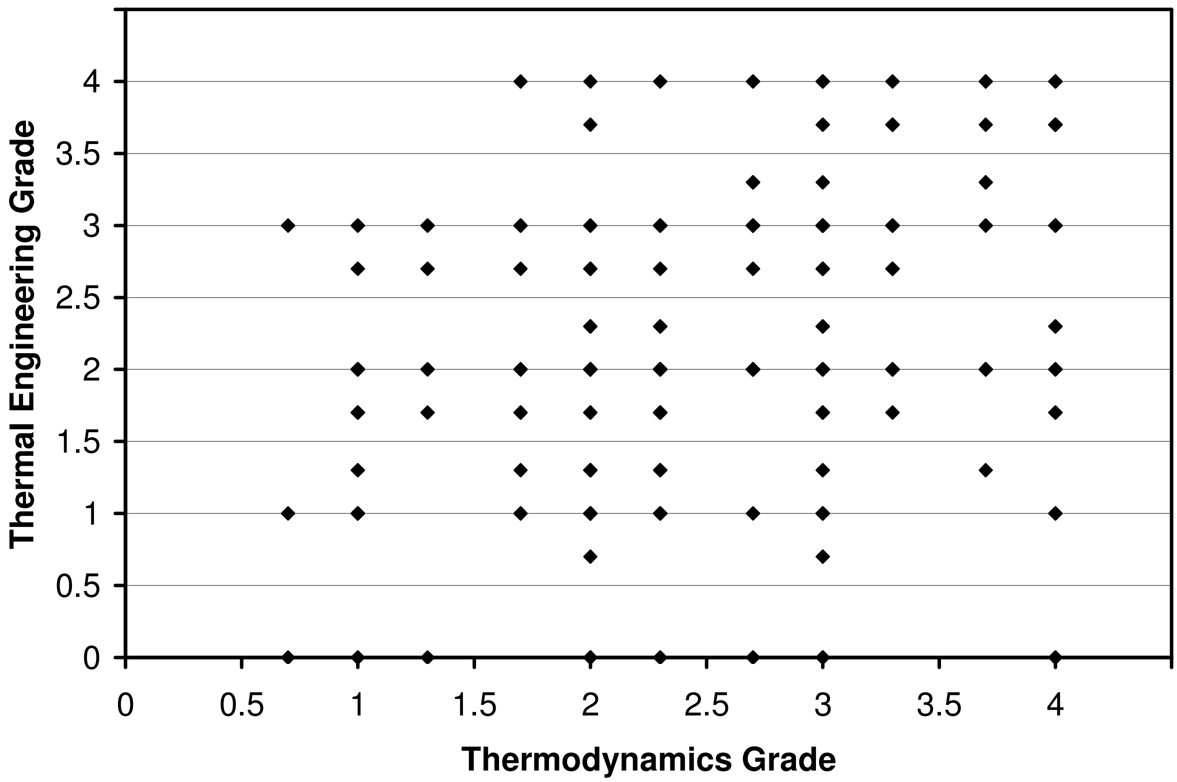

Figure 3. Comparison of Thermal Engineering Grades with Thermodynamics Grades (the prerequisite) by Student

Looking into the data more closely indicates many reasons for this including variability in faculty grading in Thermodynamics and variability in student achievement from term to term and class to class. Average GPAs for the students involved in this comparison given by the 9 Thermodynamics faculty members were: 2.1, 2.3, 2.4, 2.4, 2.5, 2.5, 2.7, 3.1, and 3.3. This essentially makes it impossible to use prerequisite grade as a predictor success in later, related courses. Normalizing grades in the prerequisite by average grade given per faculty member did not dramatically change this figure, indicating that student performance (again, as measured by grade earned) simply varies a lot from course to course. With the data available, we are unable to hypothesize further regarding reasons for this variability, but it is worthy of future investigation.

The one exception to the variability noted above could be when the same faculty member teaches both the prerequisite and the follow-up course. We searched through the database and found a recent example where a single faculty member taught both our Engineering Dynamics and Engineering Statics (the prerequisite) courses in sequential quarters. Performing a similar analysis as shown in Figure 3 but for these two courses taught by the same faculty member resulted in a figure virtually identical to Figure 3. The R^2 value increased slightly to 0.11, but even when taught by the same faculty member, performance in a prerequisite does not appear to be a good predictor of follow-up success. Because our program allows students to choose when to take classes, there will be some inherent bias in this data; if a particular student did not like this faculty member for Statics, they could choose to wait to take Dynamics with someone else. Given the very low correlation, we do not feel this bias in the data would alter our conclusion. An area for further study will be to use our PeopleSoft database to understand if students who choose to delay taking a follow-up course experience a loss of retention due to the delay.

We were also curious if student success in early ME courses could be used to predict overall cumulative GPA at graduation. To begin to answer this question, we extended our previous analysis to compare Engineering Dynamics grade with GPA at graduation for a random sample of students and found a better correlation ($R^2 = 0.4$), but this is still too low to be useful in predicting student success. For instance, many students who failed Engineering Dynamics during the particular quarter we investigated (Fall 1999) went on to graduate with a quite respectable GPA of 3.0 or higher. We caution those wishing to perform this type of analysis because it was dramatically more time intensive than those previously discussed. Analyzing information by course means that several clicks of the mouse retrieve the results for ~35 students. As soon as you want graduation data, you need to access each student's record individually which can quickly become overwhelming given the number of students enrolled in these courses. Automated queries would certainly be an advantage in addressing this sort of question. All in all, our analysis for Thermodynamics/Thermal Systems and Statics/Dynamics indicate that grades in a prerequisite are, at best, a weak predictor of future success. Others have found that much greater level of background information is necessary to develop predictors of future success (Felder et al., 1993).

As members of a large mechanical engineering department (~25 tenured or tenure-track faculty members), we were also interested in understanding how much an individual faculty member can influence future student success. It can be fairly controversial to identify stronger or weaker members of one's department, but we wanted to see if on-campus databases could provide this ability. As previously discussed, there is significant variability in student performance from term to term, so we were not confident of the database's ability to compare faculty members' performance. There are many ways to address this question, but the easiest method for us was to use data already discussed.

Nine different faculty members taught Thermodynamics in 2004-2006, and we wanted to see if success in the subsequent course, Thermal Engineering, related to which Thermodynamics faculty member taught a particular student. Of these nine faculty, all had approximately 20 students or more involved in the analysis shown in Figure 3 except one, who we excluded from this comparison because of too small of a sample size. Table 1 (next page) presents students grouped by their Thermodynamics faculty member and then lists their average grades in Thermal Engineering. Statistically, it is not possible to discern differences in instructor performance with the exception of Professor A who does appear to preparing their students less well, particularly when compared to Professor H.

Table 1. Average GPAs for Students in Thermal Engineering, Grouped by Which Faculty Member the Students had for Thermodynamics (the prerequisite).

	GPA	95% CI	n =
Prof. A	1.4	0.5	19
Prof. B	1.9	0.4	28
Prof. C	2.1	0.3	43
Prof. D	2.1	0.4	22
Prof. E	2.2	0.3	17
Prof. F	2.2	0.2	67
Prof. G	2.3	0.3	16
Prof. H	2.5	0.5	27

To further understand this, we decided to compare just these two faculty members, both of which are tenured. Figure 4 compares the grades given by Professor A and Professor H in Thermodynamics by student with the grades those students received in Thermal Engineering. Also, we have included a line of slope = 1 to facilitate comparison. Looking first at Professor A, note that none of their data points appear above the line of slope = 1. Compare this to Professor H, whose data points generally fall above this line. These observations do not address teaching quality, but simply grading rigor with Professor A being more generous than the two faculty members who taught Thermal Engineering and Professor H being generally less generous.

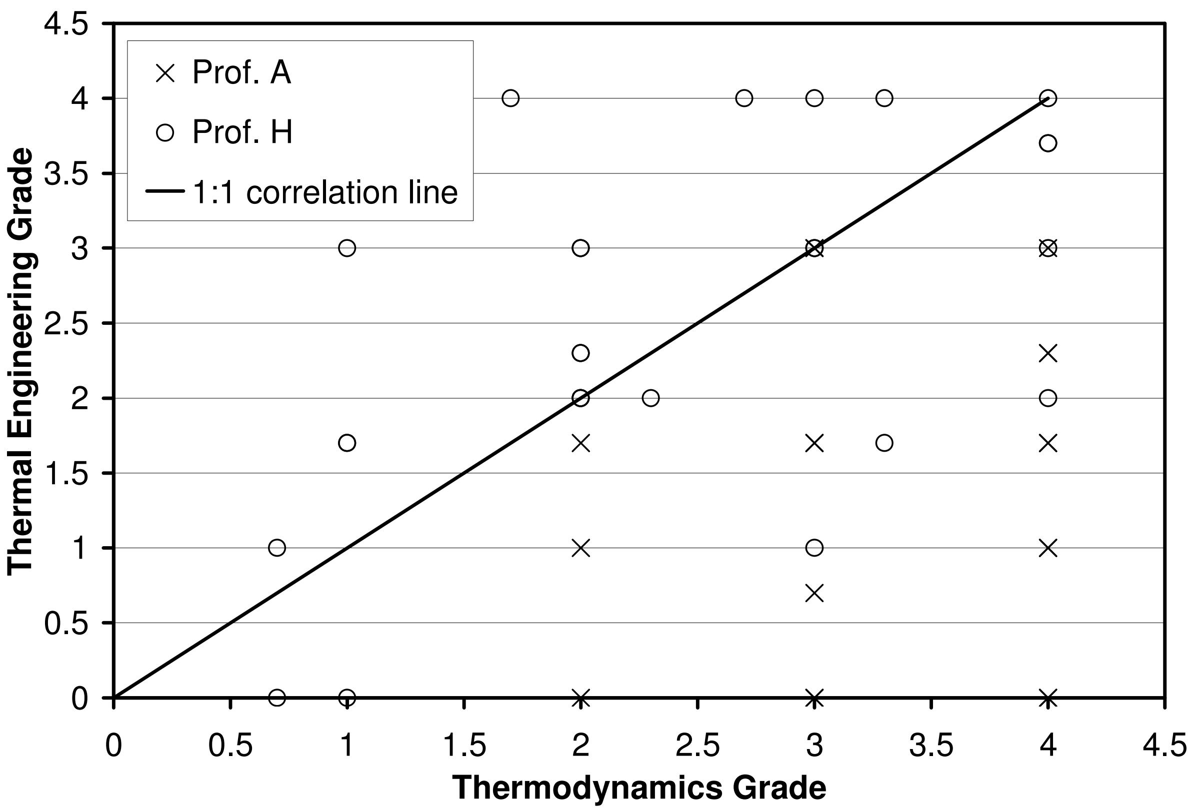

Figure 4. Comparison of Thermal Engineering Grades with Thermodynamics Grades (the prerequisite) by Student, For Two Specific Thermodynamics Professors.

To directly address teaching quality, note the horizontal line representing a GPA of 2.0 (a C grade). It would be expected that students who were well prepared in the prerequisite would generally fall above this value, and students with poorer preparation in the prerequisite would generally be below this value. Seven of the 27 students (~25%) who had Professor H earned below a C in Thermal Engineering. At first this may appear not particularly noteworthy, but compare this to 14 of 19 students (~75%) earned below a C in Thermal Engineering for Professor A. It is clear that some of our faculty are preparing their students better than others for future success.

Conclusions

The authors have presented an alternative method of indirect assessment of engineering education programs which utilizes existing on-campus databases to assess student, faculty, and department performance. Presented here is a pilot-study which we hope will pique the interest of others. Based on this effort, we feel that additional investigation, possibly with automated queries, is warranted. A few results specific to the Mechanical Engineering Department at Cal Poly are presented, but the scope of questions which can be answered with this approach is extremely broad. While the methodology is not revolutionary, we hope to have provided the motivation for other individuals to take advantage of the incredible resource that on-campus student databases provide.

Acknowledgements

The authors would like to thank Dr. James Widmann, Vera Vasquez, and Sondra Gorman, all of Cal Poly, for their assistance with this project.

References

Astin, A.W., *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education*, ACE/Macmillan, New York, 1991.

Ewell, P.T. (Ed.), "Enhancing Information Use in Decision Making," New Directions for Institutional Research, 1989.

Ewell, P.T., "National Trends in Assessing Student Learning," *Journal of Engineering Education*, April 1998, pp. 107-113.

Felder, R.M., Forrest, K.D., Baker-Ward, L., Dietz, E.J., and Mohr, P.H., "A Longitudinal Study of Engineering Student Performance and Retention: I. Success and Failure in the Introductory Course," *Journal of Engineering Education*, January 1993, pp. 15-21.

Levitt, S.D. and Dubner, S.J., Freakonomics: *A Rogue Economist Explores the Hidden Side of Everything*, William Morrow, 2005.

Olds, B.M., Moskal, B.M., and Miller, R.L., "Assessment in Engineering Education: Evolution, Approaches, and Future Collaborations," *Journal of Engineering Education*, January 2006, pp. 13-25.

Pelligrino, J.W., Chudowsky, N., and Glaser, R., (Eds.), *Knowing What Students Know: The Science and Design of Educational Assessment*, The National Academy Press, 2001.

Rogers, M. "EC2000 and Measurement: How Much Precision is Enough," *Journal of Engineering Education*, April 2000, pp. 161-165.

Shaeiwitz, J.A., "Outcomes Assessment in Engineering Education," *Journal of Engineering Education*, July 1996, pp. 239-246.