AC 2009-394: IS STUDENT PERFORMANCE DECLINING? A LOOK AT TWENTY-FIVE YEARS OF DATA

Kathy Schmidt, University of Texas, Austin

KATHY J. SCHMIDT is the Director of the Faculty Innovation Center for the Cockrell School of Engineering at the University of Texas at Austin. In this position, she promotes the School's commitment to finding ways to enrich teaching and learning. She works in all aspects of education including design and development, faculty training, learner support, and evaluation. Contact k.schmidt@mail.utexas.edu

Mark Maughmer, Penn State University

MARK D. MAUGHMER is a professor of Aerospace Engineering at the Pennsylvania State University. When he is not teaching, he is involved in the design, analysis, and wind-tunnel testing of airfoils. Contact: mdm@psu.edu

Is Student Performance Declining? A Look at Twenty-Five Years of Data

Abstract

Much of the emphasis on today's teaching is on the process (how one acquires information) rather than on the product (the information acquired). Yet for those of us who have been in the classroom for years we know the product is also valued. In looking back over one engineering professor's twenty-five years of teaching and testing, we observed a decline in the amount of material covered and in the level of performance in students' grades. Recent research on undergraduates found that only about a third of fulltime students are spending 40 hours a week attending class and studying, and this 7% decline from the previous decade could be a factor in declining performance. Additionally research indicates that today's students are spending less time trying to make connections between the facts and ideas to practical applications. In this paper, we look at grading trends and amount of content covered in two aerospace engineering courses in an attempt to quantify changes and to begin a conversation on how to address the need to engage students in the diligence of learning so that they can leave our classrooms schooled with enough depth of knowledge.

Introduction

Funding, research activities, and autonomy help to make American institutions of higher education prestigious. Furthermore, the United States leads the way in access to higher education with a rate of growth since 1947 that is three times faster that of the population.[1] Given that these institutions have resources along with large numbers of students to select from, you would expect increasing numbers of college graduates. Yet there is evidence to the contrary. For although the US is one of the leaders when it comes to college participation, it is in the bottom half of college completion.[2] Some evidence suggests that graduation rates are declining because those entering college are less prepared than their predecessors.[3] Broader views stress that the issue of assessing quality in higher education is a complex one and needs multiple measures. Reports such as "Measuring Up 2008: The State-By-State Report Card for Higher Education" offer systematic data on these six categories: preparation, participation, affordability, completion, benefits, and learning.[4] Other sources, such as the book, *Declining by Degrees*, paint a broad picture that suggests, "higher education, long viewed as the crown jewel of American education, is tarnished."[5]

Higher education is under scrutiny and currently the most commonly used outcome indicator is retention rates. Although retention is telling, we should note that "retention and graduation rates are a very primitive outcome measure: They beg the question of whether, and what, students have actually learned."[6] One way to unearth what students are learning is to study, at a macro-level, what goes on in a classroom or a set of classrooms. Professors deal first hand with students' readiness and willingness to learn and are able to capture student performance data. Perhaps there is hesitation though to share decreasing performance for as Trimble notes, "no college would embarrass itself by showing that the performance of its students was declining."[3] In fact, college grade point averages are at a high and there are many reports that grades are

inflated ("an increase in grade point average without a concomitant increase in achievement"[7]). In a review of the literature on grade inflation, Boertz contends, "…the literature demonstrates that the improvement is not incongruous with a rise in faculty development programs and increased varieties of student support services."[8] Grades and students' ability remain a contentious topic and not always seen as an accurate assessment of student performance. Grades, however, are one way to measure student learning.

Classroom grades generally combine several measures such as homework, projects, participation, quizzes, and examinations. The effectiveness of these measures is often studied with the most emphasis placed on the validity and reliability of the statistical reporting. Determining validity is difficult and since most professors who create these measures are not testing experts, a practical approach is to use content validity (content represents an adequate sampling of what was taught). The difficulty of the questions is of concern as well, for as Stenner pointed out, "If you don't know why this question is harder than that one, then you don't know what you are measuring."[9] By looking at the difficulty of exam questions, you should be able to develop questions of higher content validity and to effectively target different levels of difficulty.

In this paper, we will look closely at one veteran aerospace professor's exams and resulting grades. For the past twenty years, he has stored student exams and results in two aerospace courses. In addition to his questioning overall student learning, he is finding that that it takes more and more time to cover material. By developing an approach to evaluate the level of difficulty of these courses final exams, we are able to estimate test item difficulty and to present time trends of the grades versus year and see that the distribution has changed. We realize that grades and grade reporting practices do not hold any answers in themselves, but they do illustrate the progression of performance. Our intention is to initiate a dialogue on the dilemma of declining achievement.

Background

Courses

The two courses focused on in this study have been in the aerospace engineering curriculum at Penn State University (PSU) for a number of years and offer fundamental content. One is a required junior-level course, Introduction to Aeronautics, with a required prerequisite of the first aerodynamics class and an aerospace analysis class. The other is a senior-level technical elective, Theoretical Aerodynamics, and its prerequisite is the Introduction to Aeronautics course. The junior-level course introduces students to the basic concepts of aeronautics by covering the estimation of the forces of flight and how these forces are used to predict the performance of atmospheric flight vehicles. Students who successfully complete this course should be able to:

1. perform basic computations of aerodynamic forces and moments acting on an aircraft in flight;
2. perform basic computations of propulsive forces and performance;
3. perform basic stability and control computations; and
4. perform basic performance calculations for the overall air vehicle.

Students' grades are based on two one-hour examinations, each worth 25% of the overall course grade. The final examination accounts for 30% of the grade, while quizzes, homework, and projects account for 20%.

The senior-level course, Theoretical Aerodynamics, gives students an appreciation for the practical benefits of applying classical theoretical methods to the analysis and design of airfoils and wings. This class presents complex analytic function theory and how it is used, along with boundary-layer theory, to predict forces and moments on aerodynamic bodies. Students who successfully complete this course will be able to:

1. apply classical theoretical methods to the analysis and design of airfoils and wings.
2. use methods based on complex analytic functions, along with boundary-layer theory, to predict forces and moments on aerodynamic bodies.
3. approach aerodynamic design problems and use theoretical and computational tools to work toward an optimum solution.

Grades are calculated from seven homework assignments, including one or more computer project. The homework assignments were concerned with providing practice and reinforcing the desired objectives, and more specifically, with the items listed in the course outline provided in the syllabus. In addition, there were eight short quizzes during the semester, two one-hour examinations, and a final.

Students

These are students at a large public institution. Over the last twenty five years, enrollment in aerospace engineering at PSU has fluctuated along with national trends. For the purposes of this study, we have sampled classes of students from 1984 to 2007. Table 1 provides the number of students enrolled in the introduction to aeronautics course since the spring of 1984. In the earlier years, large enrollments required this course to be taught in two sections by different instructors. Thus, in the early years, the enrollments listed only represented only half of the students enrolled. In later years, enrollment controls has limited the number of students in the major.

Table 1 *Student Enrollments in Introduction to Aeronautics*

Year	# of Students
1984	77
1985	65
1989	125
1991	108
2000	54
2003	72
2008	95

There have been fewer students in the theoretical aerodynamics course because it is an elective but, but the apparent decreased enrollments over through the years roughly represents a constant

percentage of the total number of students enrolled in the major. The number of students enrolled in this course for the years considered is provided in Table 2.

Table 2 *Student Enrollments in Theoretical Aerodynamics*

Year	# of Students
1987	44
1988	42
1991	32
2006	13
2007	16

The Curriculum

Course content has not changed significantly in the years that these courses have been taught. The professor continues to use his original lecture notes, although they have been reviewed and refined throughout the years. One way to measure the amount of content covered has been by looking at how many pages of the lectures notes were used. When the aeronautics course was first taught, 202 pages of lecture notes were presented, while in 2008, only 134 pages of lecture notes were covered. The same situation exists for the senior-level elective.

The theoretical aerodynamics course has been taught and refined for over twenty years by this professor, and the targeted outcomes for the course haven't changed much. The homework and test results indicate that the students are learning the material covered satisfactorily. Perhaps of most significance, the student understanding of the design process seems to benefit a great deal from the open-ended design assignments. Unfortunately, with each passing semester, he has found that it takes more and more time to adequately cover the mathematical concepts and tools that are later employed in the classical aerodynamic theory. This problem was exacerbated the last year it was taught in that the students also came to the course poorly prepared in the basic understanding of applied aerodynamics. Thus, these concepts had to be "reviewed" before going on to the topics outlined in the syllabus for this course. Consequently, the number of advanced aerodynamic topics was less than it had been in previous years.

Grading of Final Exams

Given that there is less coverage of material, the final exams are not as long as they were in earlier years. Fundamental questions have remained on the exam, however, and have been reused throughout the years. The students do not get a copy of their final exam and therefore these exam questions can be reused. The exams are intended to take the allotted time (up to 2 hours) and students can bring in a crib sheet. Many test items require more than one ability to obtain a correct response and these exams are rigorous, although different levels of understanding are tested. Roughly, a third of the test deals with the basic understanding of the material, as third with the ability to apply the methods covered in the course to perform basic engineering analysis, and a final third on the deeper insights required to synergize the course material to address open-ended design-type problems.

Letter grades were computed on a curve. A summary of the numerical scores recorded for a number of year's worth of exams are presented in Figures 3 and 4. In order to assure fairness and consistency, established criteria is applied in the grading of these exams.

Figure 3 *Student Final Exams Results in Introduction to Aeronautics*

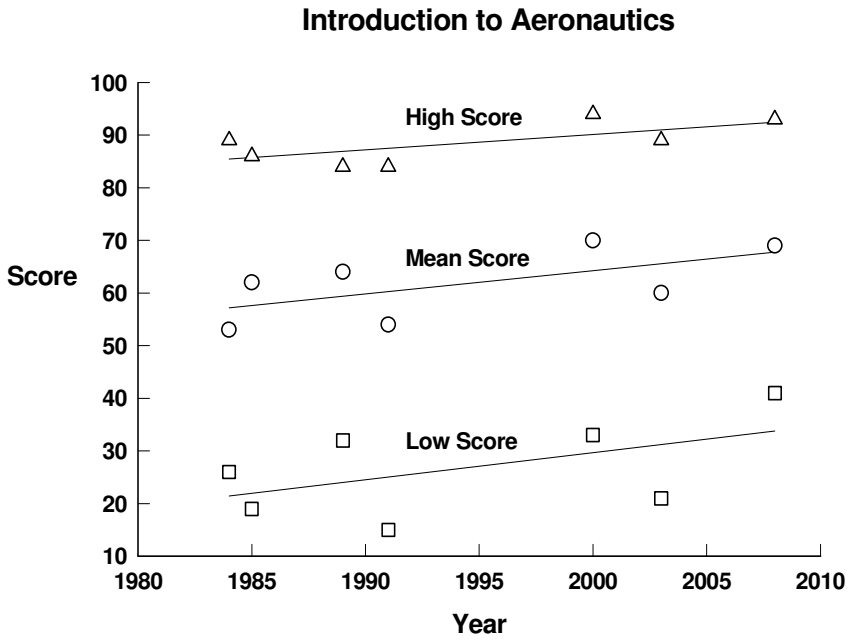
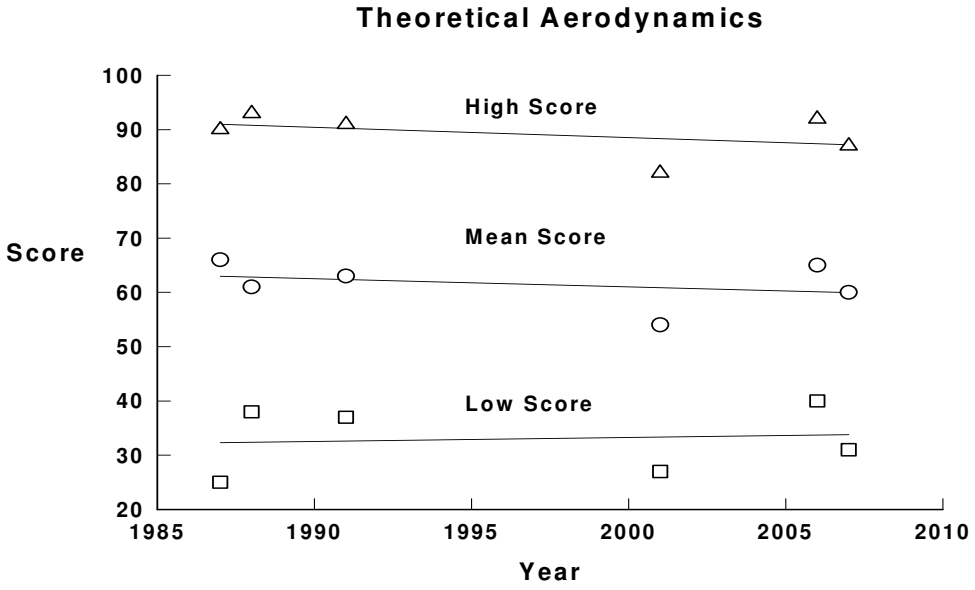


Figure 4 *Student Final Exams Results in Theoretical Aerodynamics*



Review of Exam Question Difficulty Levels

Establishing the item difficulty of each exam question was done by using a rubric. The rubric is based loosely on Bloom's Taxonomy,[10] which classifies cognitive behaviors into six categories which range from simple to complex and how much work is involved in answering and current Research has identified levels of difficulty that are as follows: concept difficulty – the intrinsic difficulty of the concept itself; process difficulty- the difficulty of cognitive operations and the demands made on cognitive resources; and question difficulty – which may be rooted in the language of questions, the presentation of questions, and the use of mark schemes in rewarding processes.[11] We took from Bloom and current research on difficulty to create the rubric illustrated in Table 2.

Table 2 *Rubric for Assessing Question Difficulty*

Difficulty	Response Effort
D1 = knowledge/comprehension	RE1 = responding such as to a multiple choice
D2 = application	RE2 = calculations such as putting numbers in an equation
D3 = procedural	RE3 = involved calculations that are procedural and often require integration
D4 = synergy or design	

The first element in establishing how "hard" an exam question is to answer is addressed by assigning a level of difficulty to it. The first level of difficulty, D1, is taken to be that one knows, for example, a definition or something only requiring rote memory. The second level, D2, considers knowing, for example, what equation to apply in a given situation, and how to use it to calculate, say, the lift force generated by a wing. The next level of difficulty, D3, involves a somewhat deeper understanding of the theoretical concepts than required by simply applying a formula. This requires, for example, the application of several concepts in a procedural manner in order to obtain an answer to a problem. This would include, for example, a calculation using thin-airfoil theory or employing some of the theoretical concepts embodied in finite-wing theory. The final level of difficulty, D4, deals with the application of ones insights and knowledge to an open-ended, design-type problem.

As one would expect, testing at the highest level of difficulty does not occur very often in the junior-level introductory course, but with much more frequency in the senior-level elective course.

In addition to the level of difficulty experienced in taking an exam, the amount of actual effort required to answer each given question is a factor in the difficulty of the exam. This is addressed by assigning a response effort to each question. The lowest level of response effort, RE1, is considered to be the effort required to give a one-word answer or choose the number of a multiple-choice type question. The next level of effort, RE2, requires performing a calculation or two, or perhaps providing a sentence or two defining an idea or concept. Finally, the highest level of effort, RE3, entails performing sequential calculations, and/or perhaps use previously learned skills such as evaluating an integral or making use of trigonometry.

Results of Difficulty Level of Final Exam Questions
To obtain an overall level of difficulty/effort for an exam, we assigned a difficulty and response effort to each exam problem. These two scores were then simply multiplied together for each problem and summed for all the problems on the exam. The result is an overall numerical assessment of how "hard" an exam is. Over the years considered, it was found the final exams in the Introduction to Aeronautics course decreased from 122 in 1984 to 87 in 2007. In the Theoretical Aerodynamics course, the overall "hardness" dropped for 150 in 1987 to 98 in 2007.

Discussion

This one study is not indicative that a declining academic performance of college students is the case. We note, however, there are other indicators out there. What are the factors that contribute to this dilemma? Reports suggest a myriad of issues, including for example, declining faculty expectations, administrative policies (i.e., relaxed admission policies), changing student attitudes, changing student demographics, inadequate text books, and the slipping of admission test scores. Furthermore, reading is on a decline and as Americans spend less time reading their reading comprehension skills are eroding. Certainly there are many ramifications in higher education to a less literate public since the number of books in a home is a significant predictor of academic achievement.[12] Our students may be entering higher education underprepared.

There is no doubt that current students have some talents and skills that exceed their predecessors. They are, for example, extremely proficient at handling and manipulating large sets of data. They are also exceptionally good at using information from texts and documents. Yet they are not being exposed to as much technical depth. In engineering, we have to ask is this a good thing? If it isn't, we have to ask "why isn't somebody doing something about it?"

It may be that these students, bombarded by a world of technology, may need directed help in staying focused. We might also work at sharing with them more on how to learn and how to study. They often enter higher education with unrealistic expectations about what it takes to be a successful college student and while there are efforts underway at many universities, individual instructors, may need to spend more time on teaching thinking skills. The classroom is a powerful forum for modeling how to be and think like an engineer and we will need to make concerted efforts to show our students what they can and should do.

Conclusions

In this case study, we found that there was a significant decrease in the difficulty level of the exams given to aerospace engineering students. Often blame is placed on the students' lack of prerequisite knowledge and while that is a contributing factor, we believe the problem is much deeper. At question is what can and should be done to rectify this situation? While setting a high standard is paramount to student achievement, we realize that you still have to teach so that students can succeed. How can you appropriately challenge and prepare your student without relaxing course content and requirements? Perhaps identifying there is a problem is the first step in working towards a solution.

References

1. Based on a population of 144,126,071 in 1947 and 262,764,948 in 1995. US Population from 1900, http://www.nas.org/polReports.cfm?Doc_Id=85. Accessed January 21, 2009.
2. Wolff, R. (2007). *The decline of public higher education.* http://ricardo.ecn.wfu.edu/~cottrell/ope/archive/0702/0075.html . Accessed January 21, 2009.
3. Trimble, S. (1986). Declining student performance in college geography and the down-writing of texts, *The Professional Geographer,* 38(3), pp. 270-273.
4. The National Center for Public Policy and Higher Education. (2008). Measuring up2008. http://measuringup2008.highereducation.org/index.php. Accessed January 23, 2009.
5. Hersch, R. H. & Merrow, J. (2005). *Declining by degrees; Higher education at risk*. New York, NY: Palgrave Macmillan.
6. Bennett, D. (2001). Assessing quality in higher education. *Liberal Education*, 87(2). http://www.aacu.org/liberaleducation/le-sp01/le-sp01bennett2.cfm. Accessed January 23, 2009.
7. Potter, W., and Nyman, M. (2001). Be careful what you wish for: Analysis of grading trends at a small liberal arts college. *College and University: The Journal of the American Association of Collegiate Registrars* 76 (4).
8. Boretz, E. (2004). Grade inflation and the myth of student consumerism. *College Teaching*, 52(2), 42-5.
9. Fisher-Hoch, H., & Hughes, S. (1996). What makes mathematics exam questions difficult? Research and Evaluation, University of Cambridge Local Examinations Syndicate, http://www.cambridgeassessment.org.uk/ca/digitalAssets/113963_What_Makes_Mathematics_Examination_Questions_Difficult.pdf, Accessed January 23, 2009.
10. Bloom, B. S. (1956). *Taxonomy of Educational Objectives. Book 1, Cognitive domain*. , Longman, New York.
11. Pollitt, A., Hutchinson, C., Entwhistle, N., & de Luca, C. (1985). *What makes examination questions difficult*? Scottish Academic Press.
12. National Endowment for the Arts. (2007). To read or not to read: A question of national consequence, http://www.arts.gov/research/ToRead.pdf. Accessed March 12, 2009.