# ARE OUR STUDENTS BETTER NOW?

W STEPHEN WILSON

Abstract. The author gave his fall 2006 Calculus I class the final exam from his fall 1989 Calculus I class. The scores were significantly lower in 2006. Although the University has changed dramatically in those 17 years, much is the same. The SAT mathematics test does not detect this decline. The conjectured culprits are the general decline in K-12 mathematics education and the use of calculators on the SAT mathematics test.

January 22, 2007

## 1. introduction

Professors are constantly asked if their students are better or worse today than in the past. This paper answers that question for one group of students.

For my fall 2006 *Calculus I for the Biological and Social Sciences* course I administered the same final exam used for the course in the fall of 1989. The SAT mathematics (SATM) scores of the two classes were nearly identical and the classes were approximately the same percentage of the Arts and Sciences freshmen. The 2006 class had significantly lower exam scores.

This is not a traditional research study in mathematics education. The value of this study is probably in the rarity of the data, which compares one generation to another.

## 2. basic comparisons

Johns Hopkins University (JHU) changed dramatically in the 17 years between 1989 and 2006. The freshmen class has grown from 831 to 1,211, a 46% increase. In 1989 JHU had 5,585 applicants for the freshmen class; in 2006 there were 13,863 applicants, a 148% increase. The acceptance rates were approximately 49% and 27% respectively and the percentage of those who enrolled after being accepted were 30% and 32% respectively. JHU is now much more selective than it was in 1989.

The 1989 and 2006 classes are surprisingly similar. I taught *Calculus I for the Biological and Social Sciences* in the fall of both years. In the fall of 1989, 147 Arts and Sciences freshmen took the course. This was 23.7% of the Arts and Sciences freshmen class of 621. In the fall of 2006, 180 of 792 Arts and Sciences freshmen, 22.7% of the class, took the course.

The average SATM score for the 1989 Calculus I class was 662.6 with a standard deviation of 6.8. For the 2006 Calculus I class it was 664.9, with a standard deviation of 6.3. In the mid-1990s, SATM scores were "recentered," [Rec07]. After

Thanks to Saul Wilson for his careful and detailed editing.

recentering[1] the 1989 class's SATM scores, the new average was 662.3, with a standard deviation of 6.5. I chose 1989 to compare with because it was the only year I both taught the course and had the SATM scores for.

For statistical analysis I rely on a colleague, Daniel Q. Naiman, Professor and Chair of the Department of Applied Mathematics and Statistics at JHU. He performed the Wilcoxon test on the SATM scores for the two years yielding a p-value of .6. From this, it can be concluded that there is no statistically significant difference in the SATM score distributions between the two years.

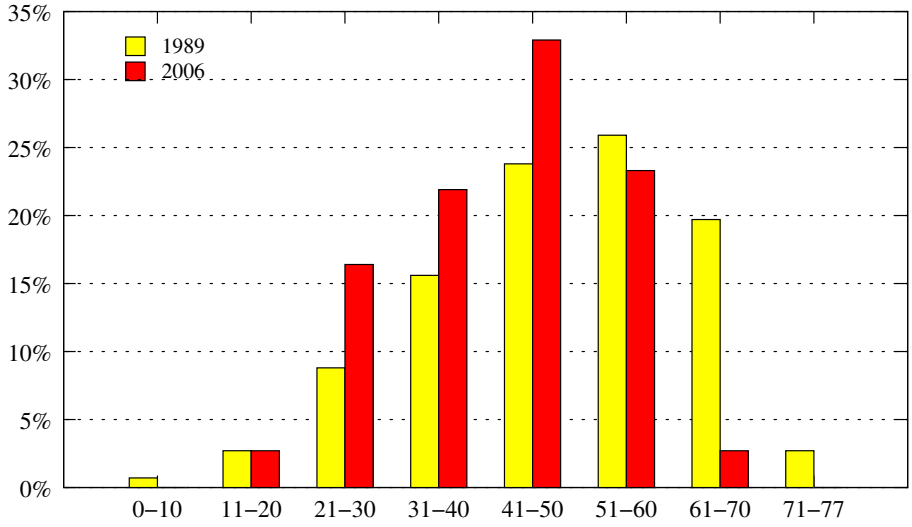
## 3. THE EXPERIMENT

The 2006 Calculus I class took the same 77-point final exam as the 1989 class. The content of the Calculus I course has not changed, and, mathematically, using the old exam was completely appropriate.

The scores on the final exam were markedly different. The average of the 1989 scores was 48.4, with a standard deviation of 14.4, while the 2006 class average was 42.5, with a standard deviation of 11.3. The 5.9 point decrease in the average is a 12.2% decline. Daniel Naiman also ran the Wilcoxon test of significance on these two distributions and found a p-value of .001 for the two-sided test.

A graphical image of the difference in raw scores, produced with the help of Matthew Peterson of the MIND Institute and Derek Bruening's front end to gnuplot for bar graphs, is provided in Figure 1.

Figure 1: Comparison of raw scores on final exam




How significant is this change educationally? Contemplate a similar drop in SATM scores. SATM scores range from 200 to 800. If there had been a 12.2% drop over the 17 years from the recentered SATM score of 662.3, the 2006 class would have an average SATM score of 605.9 ($= 662.3 - .122 \times (662.3 - 200)$). This would

[1]It is part of SAT mythology that all scores increased when they were recentered. The reality of the mathematics recentering is that scores in the 600-650 and 720-740 ranges, as well as the top score of 800, remained the same; scores below 600 and in the 750-790 range increased; those in the 660-710 range decreased. Many of the 1989 students were in the latter range.

be a shocking development at JHU, even if the Calculus knowledge of the students was the same or superior.

However, the SATM scores have not changed, but the content knowledge at the end of the first semester is down one-eighth. If this was universal among our students, it would be akin to missing a full semester during each student's undergraduate career.

Daniel Naiman suggested that I compute what grades would be on the old grading scheme. The results are in Table 1. Graphically, this is given in Figures 2 and 3. This approach firmly establishes that the drop in scores represents a significantly different level of achievement.

| Percentage of grades under different grading schemes | | | | | |
|---|---|---|---|---|---|
| 1989 Grading Scheme | | | 2006 Grading Scheme | | |
| Grades | 1989 class | 2006 class | Grades | 1989 class | 2006 class |
| A | 26.5 | 5.5 | A | 52.4 | 31.5 |
| B | 23.1 | 20.5 | B | 25.9 | 37.0 |
| C | 38.1 | 54.8 | C | 19.0 | 30.1 |
| D | 9.5 | 16.4 | D | 2.0 | 1.4 |
| F | 2.7 | 2.7 | F | .7 | 0 |

TABLE 1

## 4. INTERPRETATION

It must be confronted that the 2006 students did not do as well as the 1989 students, no matter how one tries to explain it. An easy explanation is to assume that this is the result of a slowly degenerating mathematics professor. I am not inclined to look favorably upon that explanation. Aside from my belief that I get better at teaching every successive year, I received a teaching award, *The Johns Hopkins University Homewood Student Council Award for Excellence in Teaching*, in 2000 - closer to 2006 than 1989. My student course evaluations have remained consistently high (although the results for this class will not be available for months).

If the percentage of Arts and Sciences freshmen taking Calculus had increased, then we might be encountering weaker students who, in 1989, would not have taken Calculus at all. Since the percentage in Calculus I is the same, this explanation would require an increased percentage of freshmen taking Calculus II. However, the corresponding fall semester percentages for Calculus II are 11.1% for 1989 and 11.4% for 2006.

I think it is unlikely that the phenomenon we are seeing is a result of something happening at JHU once students arrive. I am inclined to conclude that these 2006 students are not as well prepared as the corresponding group was in 1989, despite there being many more American high school graduates now and significantly more competition to get into JHU today than ever before.

Figure 2: Percentage of letter grades on the 1989 grading scale

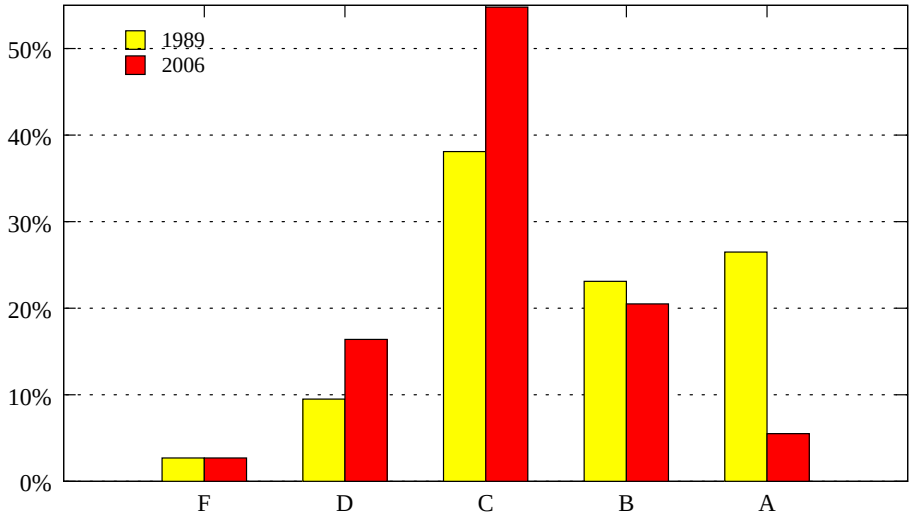
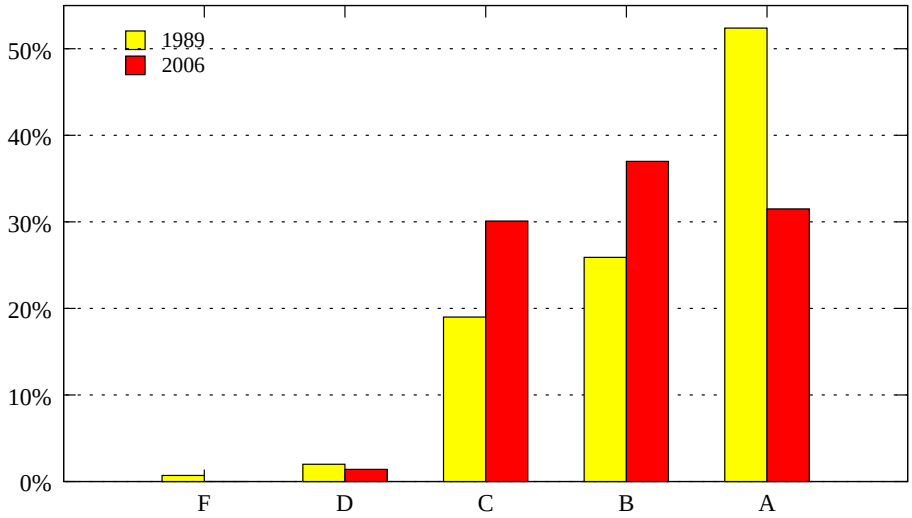



Figure 3: Percentage of letter grades on the 2006 grading scale




## 5. THE FINAL EXAM

The final exam, its distribution, and the midterm exams are on-line [Cal07a]. I give what students consider to be hard exams. Of the 77 points on the exam there are 11 one-point questions and 33 two-point questions. Although some of the problems are long, the number of points per problem is kept to a minimum so as to restrict the graders in the use of independent judgment. Consequently, grading practices likely do not account for the differences in grades from 1989 to 2006.

## 6. CONCLUSIONS AND RECOMMENDATIONS

The similarities between these Calculus I classes 17 years apart are striking. In addition to having the same teacher, the stability of syllabus, class size, and SATM

scores is surely unusual. It appears that this 2006 group of JHU students comes to the University significantly less well prepared than their counterparts in years past. This phenomenon is probably shared with many other universities.

Nineteen eighty-nine is, in mathematics education, indelibly tied to the National Council of Teachers of Mathematics' publication, *Curriculum and Evaluation Standards for School Mathematics* (1989), which downplayed pencil and paper computations and strongly suggested that calculators play an important role in K-12 mathematics education. My 2006 students would have been about two years old at the time of this very influential publication, and it could easily have affected the mathematical education many of them received. Certainly, one possibility is that mathematics preparation is down across the country, thus limiting the pool of well prepared college applicants.

There is nothing that universities can do to correct the lack of preparation of their applicants. However, it is difficult to believe that there are not enough students to fill our classes with 1989 quality students. One of the major gate-keepers, the SATM test, is oblivious to this significant shift in preparation. Universities can certainly demand a more effective SATM test.

Since 1994 the College Board has allowed the use of calculators on the test. The College Board's calculator policy, [Cal07b], states: "Every question on the SAT Reasoning Test [SATM] can be solved without a calculator; but you will gain an advantage by using a calculator with which you are familiar." I conjecture that it is precisely this gained "advantage" that causes the SATM to fail universities in the admissions process. This conjecture is consistent with the 2002 JHU study, [WN04], that found that students for whom "in K-12, calculator usage was emphasized and encouraged" had lower mathematics grades in the large service courses. As it stands, universities have no way of rejecting applicants who do not know arithmetic adequately for college-level mathematics.

I urge universities to join together to negotiate with the College Board for a more appropriate test or to look to an alternative test.

## References

[Cal07a] `http://www.math.jhu.edu/∼wsw/F06/CALC1`, 1/07.

[Cal07b] `http://www.collegeboard.com/student/testing/sat/testday/calc.html`, 1/07.

[Rec07] `http://www.collegeboard.com/repository/rs05_3962.pdf`, 1/07.

[WN04] W.S. Wilson and D.Q. Naiman. K-12 calculator usage and college grades. *Educational Studies in Mathematics*, 56:119–122, 2004.

Department of Mathematics, Johns Hopkins University, Baltimore, Maryland 21218
*E-mail address*: `wsw@math.jhu.edu`