

# Statistical Evidence that Web-Based Homework Helps

By L. Hirsch and C. Weibel

Since 1996, dozens of mathematics departments have introduced an internet-based homework system into their calculus curricula. One popular choice is WeBWorK, which was developed with NSF support at the University of Rochester. At Rutgers University, we were able to examine the effect of WeBWorK on student performance on the common final exam in our general calculus class. (This course excludes engineering, physics, chemistry and mathematics majors.) Because WeBWorK was implemented in only two-thirds of all sections, we could treat the course as a controlled experiment, using the non-WeBWorK sections as the control group. (We later verified that there were no significant section effects.)

The study used data from the Fall 2001 semester of the general calculus class at Rutgers. Out of over 1300 students in the class, we had complete data (SAT scores, placement scores as well as final exams, etc.) for 1175 students. Of these, 368 were in the control group and 807 students in the study group. Both groups submitted written homework assignments, but in the study group about 11 written problems per week were replaced by WeBWorK problems. The WeBWorK questions were similar for each student, but were individualized by varying the numerical parameters from student to student.

Overall, the students in the web-based sections showed a small but statistically significant improvement (of 4%) on the final exam. This improvement was also present after adjusting for placement scores, which are a measure of prior skill level. As reported by us in the January 2001 issue of the FOCUS, placement scores are the most significant predictor of success. The small difference between group means is not surprising, because the WeBWorK group contained a sub-

population of students who did not attempt many web assignments. This sub-population diluted the gains in final exam performance attributable to WeBWorK, as the analysis detailed below shows.

In fact, we found that the effectiveness of WeBWorK depended dramatically

who attempted most problems. This was also true after adjusting for the effect of prior knowledge, measured either by our placement test or by their Math SAT (also a significant predictor).

Surprisingly, the data suggests a quadratic relationship between the percentage of WeBWorK problems attempted and the final exam score, with the best-fitting curve being concave up. In other words, students who do less than 50% of the WeBWorK get less “marginal” benefit on the final (i.e., less improvement from doing one more problem) than students who do over 80% of the WeBWorK.

A second sub-population consisted of upper-class students who were taking calculus for the first time. Unlike the first-year students who placed right into calculus, these students typically took precalculus first. In WeBWorK sections, most of these students attempted between 40% and 80% of the assigned WeBWorK problems. Here there was a three letter-grade difference (from F to B) between those who did most

of the assigned WeBWorK problems and those who did only a few.

Upper-class students who were repeating Calculus did not seem to derive any benefit from attempting WeBWorK problems. We are still in the process of studying this group in an effort to understand the reason for the absence of any correlation between the amount of WeBWorK activity and final exam scores in this group. ■

*Lew Hirsch and Chuck Weibel teach at Rutgers University.*

◀ Previous
▲ Prob. List
Next ▶

*Your score on this attempt is 50%.*

(1 pt) Find the absolute maximum and the absolute minimum of the following function on [0,5].

$$f(x) = -34x^2 + 113x - 23$$

Absolute maximum value:

Absolute minimum value:

Show Correct Answers

*Sample WeBWorK problem. The score is 50% because one answer is wrong.*

upon how many of the problems were attempted. The correlation between attempts and percentage of problems solved was a remarkable .944, suggesting that once students began a problem they persisted until they had solved it. An analysis of variance showed that only 9% of the variability in WeBWorK scores could be attributed to prior skill level, even among entering students, so the number of problems attempted may be an indicator of effort rather than ability.

Among first-year students in WeBWorK sections, most of whom attempted every problem, there was a two-letter grade difference on the final (from D to B) between students who did not attempt many WeBWorK problems and those