

# Comparing Students' Performance On Research-based Conceptual Assessments And Traditional Classroom Assessments

N. Sanjay Rebello

*Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506-2601*

**Abstract.** The use of concept inventories to investigate students' learning gains is common in physics education research. However, comparatively little research has compared students' learning gains on concept inventories with other more traditional assessments in the classroom. We present a study comparing second semester calculus-based physics students' performance on traditional classroom assessments including exams and homework with learning gains on SEMCO (Survey of Electricity, Magnetism, Circuits and Optics), which was previously created by combining questions on other conceptual surveys such as CSEM and DIRECT. We report on students' performance on specific items on SEMCO and corresponding traditional classroom assessments that are based on the same topic. Our results indicate that while the overall performance on SEMCO might correlate with aggregate performance on class exams, the performance on clusters of SEMCO items that assess conceptual understanding in various topical areas does not correlate as strongly with performance on corresponding traditional exams. These results raise some potentially interesting issues on the validity and usefulness of traditional classroom assessments and conceptual assessments that are often used to measure student learning in introductory physics.

**Keywords:** physics education research, conceptual understanding, assessments

**PACS:** 01.40.Fk

## INTRODUCTION

Physics education researchers have often used conceptual inventories to assess student learning in introductory physics courses. As such, conceptual inventories provide a useful tool for both physics education researchers as well as educators who aim to assess conceptual learning by their students. One of the main advantages afforded by conceptual inventories is that they often serve as proxy for standardized assessments that allows for comparison of conceptual understanding between students in different classes and in different universities. Indeed, conceptual inventories have also been used in large scale comparisons of learning by students in reformed or interactive engagement classes versus those in traditional or lecture-based classes [1].

While conceptual inventories may have their advantages, they are often not commonly used by instructors. Rather, most instructors rely on exams that they create for assessing student learning in their class. Commonly cited reasons for the use of class exams rather than conceptual inventories for assessing student learning are that conceptual inventories do not assess problem solving which is often emphasized in the course and that the inventories span a limited range of topics that are only a small fraction of the topics covered in a semester-long course.

So, while most instructors often do not use inventories as assessments of their students' learning, inventories are commonly used to assess students' conceptual understanding by researchers. We thus decided to conduct a study that compares students' performance on inventories with their performance on typical in-class exams created by the instructor.

Specifically, we sought to answer the following research questions:

- 1) How does the pre-test score and normalized learning gain on a concept inventory correlate with aggregate student performance on all class exams?
- 2) How does the pre-test score and normalized learning gain on specific item clusters that assess topics covered on individual class exams correlate with the student performance on the corresponding class exams?

## METHODOLOGY

Most conceptual inventories focus on selected topics, such as force (FCI) [2], direct current circuits (DIRECT) [3], and electricity and magnetism (CSEM) [4], to name a few. Thus, few if any conceptual inventories address the entire range of topics covered in a typical first or second semester course. One exception is SEMCO (Survey of Electricity, Magnetism, Circuits and Optics). SEMCO was first

developed by Churukian and others [5] to assess student learning in second semester calculus-based physics. SEMCO consists of items from CSEM and DIRECT, as well as items on geometrical and physical optics from Mazur's ConcepTests [6]. Thus, SEMCO is an ideal instrument for the purpose of this study because its items encompass almost all of the topics covered in second-semester calculus-based physics.

The study was conducted in a course taught in the 'new studio' format [5]. The class has two hour-long lectures each week, followed by two two-hour long studios. The studios include group problem solving and laboratories. The 'new studio' model, when first implemented over 10 years ago had shown promising gains on conceptual learning compared to the previously used traditional format.

We administered SEMCO as a pre-test on the first day of studio and post-test on the last day of studio. For both the pre-test and the post-test, students were offered extra credit for completing the survey but no points were deducted for incorrect responses. Students were encouraged to provide thoughtful answers to the questions. In each case, students took about 30 minutes to complete the 38 item survey. It should be noted that three of the items on SEMCO ask students to rate their confidence on certain item clusters. Thus, the only responses analyzed for this study were on the remaining 35 items.

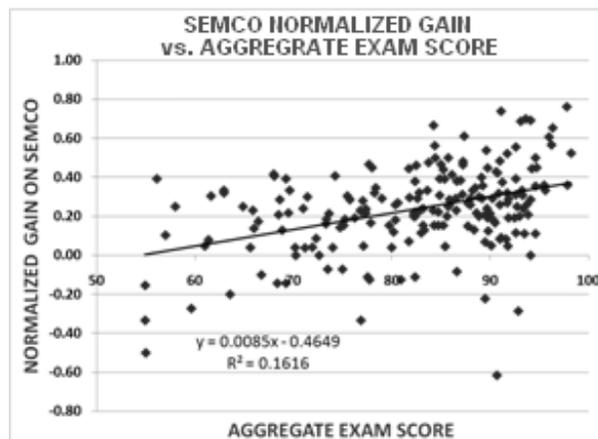
In addition to collecting data on SEMCO performance, we also collected student scores on each of the five class exams and the final exam. Each class exam was worth 100 points and together the four highest exam scores accounted for 40% of the course grade. To address our research questions, we calculated the aggregate score on the four highest class exams for each student and calculated the Pearson correlation coefficient between the SEMCO pre-score and the pre-to-post normalized gain on SEMCO calculated using the formula by Hake [1]. We also calculated the correlation between individual exam scores and the corresponding item clusters on SEMCO that assessed the same topics as the class exams.

## RESULTS & DISCUSSION

Over 88% of the 221 ( $N = 195$ ) students enrolled completed both the SEMCO pre-test and post-test. We found no significant correlation between the SEMCO pre-test score and exam performance. This result is expected because unlike the FCI [2] or FMCE [7] where the incorrect choices are distracters that are based on common student misconceptions acquired through naïve physical intuition, the incorrect choices on SEMCO items were not necessarily based on students' naïve physical intuition, since most students

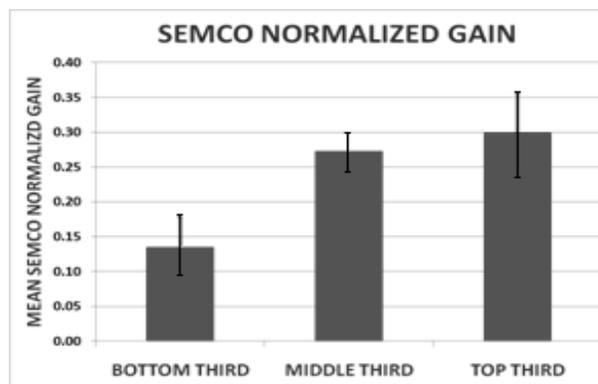
do not develop intuitions about topics such as electricity, magnetism, circuits or optics based on their everyday experiences.

We found a statistically significant, ( $p < 10^{-4}$ ) albeit small correlation ( $r = 0.41$ ) between the overall normalized gain on SEMCO and the mean aggregate exam score. Figure 1 shows a graph of the SEMCO normalized gain versus the mean aggregate exam score on the four highest scoring exams for each student.



**FIGURE 1.** A graph of normalized gain on SEMCO versus the average aggregate exam score on the four best exam scores.

Consistent with the results, an ANOVA found a statistically significant difference ( $p < 10^{-4}$ ) between SEMCO normalized gains of the students with the top, middle and bottom exam scores, as shown in Figure 2. A  $t$ -test showed no statistically significant difference between the top and middle scoring students, but did show a statistically significant difference between either of these two groups and the students with the bottom third of the exam scores.



**FIGURE 2.** Mean SEMCO normalized gains of students in the top, middle and bottom third of the aggregate exam scores.

Finally, we also calculated the correlation coefficients between the normalized gain scores on item clusters on SEMCO and corresponding exam scores that assess the same topics in the course. The results are shown in Table 1.

**TABLE 1.** Correlation coefficients between mean normalized gains of SEMCO item clusters corresponding to topics covered on each class exam

EXAM	Topics	# of SEMCO Items	r
1	Electric Field, Electric Force	6	0.33
2	Gauss's Law, Potential, Circuits	15	0.20
3	Magnetism, Faraday's Law	5	0.18
4	AC Circuits	None	--
5	Lenses, Mirrors, Polarization	5	0.25
Final	Interference, Diffraction	4	0.23

As shown in Table 1, the correlation coefficients are statistically significant ( $p < 0.01$ ), but small. They are in fact, even weaker correlations than the correlation between the overall SEMCO normalized gain scores and the average aggregate exam scores on the four best scoring exams.

Both of these results indicate that while SEMCO normalized gains might correlate weakly with scores on traditional assessments in the class, such as class exams, the correlation is even weaker when considering clusters of SEMCO items corresponding to specific topics and the corresponding exam scores.

## CONCLUSIONS

This study has demonstrated that although the overall performance on research-based multiple-choice conceptual assessments such as SEMCO might correlate weakly with the aggregate exam performance, the performance on individual item clusters correlates even more weakly with corresponding exam scores.

The reason for the weak correlation between student performance on research-based conceptual assessments such as SEMCO and traditional class assessments such as exams is most likely because they assess very different constructs. While the former focuses on conceptual understanding, traditional exams typically assess problem solving skills. Thus, to obtain a complete picture of student learning in a class, it would not suffice to use just one of those two kinds of assessments. Rather, it would be necessary to

use both research-based conceptual assessments as well as traditional exams that are typically used to assess student learning.

## REFERENCES

1. Hake, R.R., *Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses*. American Journal of Physics, 1998. **66**(1): p. 64-74.
2. Hestenes, D., M. Wells, and G. Swackhammer, *Force Concept Inventory*. The Physics Teacher, 1992. **30**: p. 141-151.
3. Engelhardt, P.V. and R.J. Beichner, *Students' understanding of direct current resistive electrical circuits*. American Journal of Physics, 2004. **72**(1): p. 98-115.
4. Maloney, D.P., et al., *Surveying students' conceptual knowledge of electricity and magnetism*. American Journal of Physics, 2001. **69**(7): p. S12-S23.
5. Sorensen, C.M., et al., *The New Studio format for instruction of introductory physics*. American Journal of Physics, 2006. **74**(12): p. 1077-1082.
6. Mazur, E., *Peer Instruction: A User's Manual*. 1997, Upper Saddle River, NJ: Prentice-Hall.
7. Thornton, R.K. and D.R. Sokoloff, *Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula*. American Journal of Physics, 1998. **66**(4): p. 338-352.