From the *Proceedings of the 2000 meeting of the American Society of Engineering Education*

Session 2380

Evaluating introductory physics classes in light of the ABET criteria: An example from the SCALE-UP Project

**Jeffery M. Saul, Duane L. Deardorff, David S. Abbott,
Rhett J. Allain, and Robert J. Beichner
North Carolina State University**

Abstract

The Student-Centered Activities for Large Enrollment University Physics (SCALE-UP) project at North Carolina State University (NCSU) is developing a curriculum to promote learning through in-class group activities in introductory physics classes up to 100 students. We are currently in Phase II of the project using a specially designed multimedia classroom for 54 students to teach the introductory physics course for engineering majors. This is an intermediate step to the full SCALE-UP classes (99 students) that will be taught in Fall 2000 when the larger classroom is completed. Both classrooms are designed to encourage students to work in groups of three, provide each group with to a laptop computer that has access to the Internet, and allow instructors to interact with each student group. Traditional lecture and laboratory are replaced with an integrated approach using active-learning cooperative group activities. The project is investigating several aspects of instruction including classroom design, classroom management, and curriculum materials. The curriculum materials include adaptation of research-based/informed activities from the literature to the SCALE-UP classroom and development of new activities. This talk will focus on the evaluation of the project, in particular, evaluating whether students are achieving the learning objectives for the curriculum. Several of the course learning objectives overlap the ABET 2000 criteria including: learning to work well in groups (teamwork), communicating effectively, being able to apply knowledge of mathematics and physics to new situations, and conducting, analyzing and interpreting experiments in addition to building a functional understanding of the course content. Evaluation methods of the SCALE-UP classes taught during the 1998-2000 school years include concept tests, individual and group exams, peer evaluation, and focus group interviews. The results show that students are building a better understanding of the main physics concepts, are more successful at solving problems, and are generally on-task and communicating well during group activities.

I. Introduction

The ABET 2000 criteria represent a radical departure in evaluating undergraduate education programs. Rather than require engineering and technology programs to conform to a set of national standards, the ABET 2000 criteria requires these programs to define their own learning objectives, use these objectives to develop measurable outcomes, evaluate the achievement of these outcomes, and use the evaluation feedback to make improvements in an iterative cycle. In addition, engineering programs must demonstrate that their graduates have achieved the 11 learning objectives shown below in Table 1.

This type of outcomes-based evaluation is not limited to major programs of study; it can also be used to evaluate and improve courses. Take for example, the calculus-based introductory physics sequence. The learning objective of this sequence is to help students build a good functional understanding of physics and develop problem-solving skills so that they can use what they learn to solve problems in new contexts. This requires students to develop multiple skills including the following:

- to be able to understand and use fundamental physics concepts,
- to know when and where specific concepts apply,
- to be able to express their functional understanding in multiple representations including graphs, diagrams, equations, and words, and
- to understand the nature of physics and how to use it effectively in and out of class.

By using evaluation and assessment methods beyond typical end-of-chapter problems (similar to those described in Section II), physics educators found that many, if not the majority, of students

Table 1: ABET 2000 Criterion 3. Program Outcomes and Assessment

Engineering programs must demonstrate that their graduates have:

1. an ability to apply knowledge of mathematics, science, and engineering
2. an ability to design and conduct experiments, as well as to analyze and interpret data
3. an ability to design a system, component, or process to meet desired needs
4. an ability to function on multi-disciplinary teams
5. an ability to identify, formulate, and solve engineering problems
6. an understanding of professional and ethical responsibility
7. an ability to communicate effectively
8. the broad education necessary to understand the impact of engineering solutions in a global and societal context
9. a recognition of the need for, and an ability to engage in life-long learning
10. a knowledge of contemporary issues
11. an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.

Each program must have an assessment process with documented results. Evidence must be given that the results are applied to the further development and improvement of the program. The assessment process must demonstrate that the outcomes important to the mission of the institution and the objectives of the program, including those listed above, are being measured. Evidence that may be used includes, but is not limited to the following: student portfolios, including design projects; nationally-normed subject content examinations; alumni surveys that document professional accomplishments and career development activities; employer surveys; and placement data of graduates.

who took introductory physics with traditional lecture and laboratory instruction had the following difficulties:[1]

- a weak grasp of basic physics concepts,
- an inability to apply what they know to new situations,
- a belief that physics is just a collection of equations and procedures that deal with very specific situations,
- a belief that physics does not have anything to do with their everyday life, and
- the failure to see physics as a process of trying to make sense out of the physical world.

Careful study of student learning in introductory physics classes has resulted in the development of several innovative introductory physics curricula that are more successful in meeting some or all of the objectives described above, including the SCALE-UP curriculum.[2] SCALE-UP is an activity-based curriculum for the calculus-based introductory physics sequence that is designed to be effective at meeting the learning objectives described above as well as being cost effective for teaching the 1000-1500 students who enroll in the sequence every semester at NCSU, 80% of whom are engineering majors.) The basic idea is that lecture and laboratory are merged into a format that stresses student learning through group activities. In addition to developing curriculum materials, SCALE-UP is also experimenting with classroom design and classroom management techniques as well. A description of the SCALE-UP project can be found in the previous paper in this session "Introduction to SCALE-UP: Student Centered Activities for Large Enrollment University Physics." A description of the group activities can be found in the paper "Promoting collaborative group in large enrollment courses" in poster session 1526.

The six course objectives for the SCALE-UP courses are:

1. Students should develop a good functional understanding of physics
2. Students should begin to develop expert-like problem solving skills
3. Students should improve their communication, interpersonal, questioning, and teamwork skills
4. Students should develop good laboratory skills including being able to design, carryout, and analyze an experiment.
5. Students should be able to use computers to look up information, take and analyze data, run simulations, and to develop mathematical models of physical situations.
6. Students should perceive the SCALE-UP classes as a positive physics learning experience.

Several of these objectives overlap strongly with the ABET program objectives in Table 1: the ability to apply scientific knowledge, developing good experimental skills, the ability to work in teams, the ability to communicate effectively, and the ability to apply the techniques, skills, and tools needed for good problem solving and professional practice.

This paper looks at our evaluation of the SCALE-UP classes to see if they are succeeding in meeting the course objectives described above. Although the evaluation methods described here are used to evaluate introductory physics classes, the methods are general enough to be used with other science and engineering courses.

II. Evaluation Methods

In developing the SCALE-UP curriculum, we design, test, modify, and revise the curriculum in a continuous cycle on the basis of classroom experience and systematic investigations of the target student population. An important part of the SCALE-UP project is to evaluate both students' reactions to the SCALE-UP class and what they are learning relative to a regular introductory course. In addition, individual activities are evaluated both for current effectiveness and with an eye to problems when we begin teaching SCALE-UP to 99 students at a time. In our evaluation of the SCALE-UP classes we use the following methods:

1. Classroom observations,
2. Diagnostic testing,
3. Surveys and department course evaluation forms
4. Portfolios of student work, and
5. Interviews with students.

A. Classroom Observations

Most instructors use observations of students to some degree to judge how their class is going. For the SCALE-UP project, classroom observations are used to gauge how well the student groups are interacting, if they are on task, how well the activity went, the quality of classroom discussions and interactions between students and faculty. Each class is video taped for later analysis. In addition, the SCALE-UP instructors keep teaching journals to record their comments on student interaction with the SCALE-UP curriculum. They also record the questions students ask as a way of gauging student understanding. Besides the instructors, other members of our physics education research and development group at NCSU act as silent observers in class taking notes on how a particular activity was carried out, how much time it took, and how students reacted to the activity.

One of the key findings of physics education research is that it is important to not just listen to what the students are saying about the course material, but to draw out the students and see what they really think.[3] Classroom observations can be very helpful in this regard. However, it is difficult to make substantial observations on more than a small fraction of a class and the observations are dependent on available opportunities. To see what is happening on a class-wide level, it is necessary to use evaluation tools like diagnostic testing and exams that can be easily delivered and evaluated for a large class.

B. Diagnostic Testing

Diagnostic testing can be used to serve two purposes. Pre-course diagnostic testing can be very useful to learn more about student skills, student knowledge, and student beliefs coming into the course. Research shows that curricula that take these factors into account tend to be more effective for helping students learn. For evaluation, pre/post diagnostic testing is used to see how well students improve over a semester. Pre/post results from the SCALE-UP classes are compared with results from the parallel traditional lecture courses at NCSU and national norms, where available, to see how well the SCALE-UP students are learning.

The diagnostic tests being used by the SCALE-UP project to evaluate instruction are listed in Table 2. These diagnostic tests are multiple-choice concept tests of which the Force Concept Inventory or FCI is a prime example.[4] The FCI is the most commonly used physics concept test in the United States today.[5] It is designed to measure students' belief in Newtonian laws of motion vs. the student's common sense beliefs. It is typical of multiple-choice concept tests in that the FCI questions that are specifically designed to trigger and identify specific common misconceptions and/or other student difficulties identified by the research literature. The distracters are taken from common incorrect student responses to open-ended versions of the questions.

In a recently published study of FCI results from over 6500 introductory physics students, Hake found that a useful figure of merit for gains in students' conceptual understanding in a class was the average fraction of the possible (normalized) gain h, where h is defined as follows,[6]

$$h = (\text{class post-test average} - \text{class pre-test average}) / (100\% - \text{class pre-test average})$$

Hake collected FCI data to see if interactive engagement/activity-based curricula were more effective for teaching Newtonian mechanics than traditional lecture methods. He found the following result:

Traditional Classes (14 classes, N = 2084 students) $h = 0.23 \pm 0.04$ (Std. Dev.)

Interactive Engagement Classes (48 classes, N = 4458 students) $h = 0.48 \pm 0.14$ (Std. Dev.)

where h is averaged over classes, not students. The average normalized gain of the PER-based classes is twice as great as the average gain for traditional lecture classes. Based on Hake's work, the normalized gain is considered the best measure of improvement in student understanding of basic concepts on pre/post diagnostic tests. Hake's result has been reproduced in a second study conducted by one of the authors (JMS) of over 2,000 students in innovative and traditional introductory physics classes at ten colleges and universities.[7] The Force and Motion Conceptual Evaluation (FMCE) is another instrument used to measure student understanding of Newtonian force and motion. It covers similar topics, but is limited to linear motion and emphasizes graphical representations. It is considered more reliable than the FCI for evaluating individual students. Both tests have been used to evaluate student learning in the first semester mechanics class.

The second semester class covers electricity, magnetism, and optics.[8] In this class we use the Conceptual Survey of Electricity and Magnetism (CSEM) and the Determining and Interpreting Resistive Electric Circuits Concept Test (DIRECT). Although the CSEM does look at student

Table 2: Diagnostic tests used to evaluate the SCALE-UP project

	Multiple-Choice Concept Tests	
FCI	Force Concept Inventory	Newtonian force & motion
FMCE	Force and Motion Conceptual Evaluation	Force & linear motion
CSEM	Conceptual Survey of Electricity and Magnetism	Non-mathematical test of electricity & magnetism
DIRECT	Determining and Interpreting Resistive Electric Circuits Concept Test	DC Circuits

understanding of electricity and magnetism concepts, it is basically a non-mathematical standard electricity and magnetism final exam. DIRECT looks at student conceptual understanding of DC resistive circuits.[9] These two diagnostics are relatively new (CSEM is still under development) and not as well established as the FCI and FMCE.

C. Department Course Evaluations

The NC State Physics department collects course evaluations from the students at the end of the semester. The course evaluation form has nineteen multiple-choice items and room for written responses to the following three questions: (1) How would you describe this course to other students? (2) What do you like best about the instruction? (3) What do you like least about the instruction? Over 80% of the students in each class completed the end-of-semester evaluation. Although both student and faculty satisfaction are important in motivating student work and presumably therefore student success, the link between satisfaction and learning is highly indirect. Indeed, students whose primary goal is a good grade may find higher satisfaction in a course that produces a good grade without improved learning, since improved learning often requires time and effort. Nevertheless, comments on course evaluations are useful for comparing student perceptions between the SCALE-UP and regular classes.

D. Portfolios of Student Work

In addition to the evaluation methods described above, SCALE-UP student exams, quizzes, lab reports, and some homework assignments are copied and archived. White board presentations are sometimes recorded with a digital camera. Examination of student responses is useful to see how well students are applying their understanding of the material and their problem-solving skills. However, while there is a tendency to focus on correct answers, it is important to pay at least equal attention to student errors. The errors are often more informative about how the students are thinking about physics. The functionality of student knowledge and skills is rather well tested by this approach since the student is being asked to produce the desired knowledge within the context of a problem and without the most common and automatic triggers. In general, the exams and quizzes tend to be more useful than lab reports and homework for seeing how individual students think, since only under these circumstances are they working alone under controlled conditions.

E. Interviews

The interview method is the most effective approach for learning what students think, either in terms of their knowledge, their skills, or their opinions on what helps them learn. The interviewer can follow up suggestive responses with more detailed and individually designed questions to probe how students think in great detail, but this evaluation method is highly time consuming. In addition to the recording time (about one half hour per student), the recordings must be transcribed and analyzed – a process that takes 4-10 hours per hour of interview.

Because this method is so time intensive, it is not practical to interview all students individually in a large class. In the SCALE-UP evaluation, we use two types of interviews, focus groups (3 to 4 students) and individual interviews. The focus group interviews are used to poll student perceptions at various times during the semester. A small number of individual interviews are used to look at more specific issues such as student epistemology or problem solving skills. Because many students share a relatively small number of opinions and difficulties, a small number of interviews can usually reveal most of the common student difficulties and comments in great detail.

III. Evaluation of Learning Objectives

A. Objective: Functional Understanding of Physics

Conceptual tests and portfolios of student work are used to determine if SCALE-UP students are developing a robust functional understanding of physics. The portfolios include copies of student responses on exams and other written assignments. The exams include problems from the regular classes, conceptual problems, and complex quantitative problems. The problems from the regular classes are used to compare problem-solving performance between the SCALE-UP and regular lecture classes.

1. Conceptual Tests

The conceptual test results from both classes, shown in Table 3, are encouraging. The SCALE-UP students in the first semester mechanics class outperformed their peers in the regular lecture class on the FCI and the FMCE. Recall that according to the study by Hake, normalized gains on the FCI from pre-test to post-test for traditional mechanics classes average 23%. The fall and spring semester SCALE-UP mechanics classes from the 1998-99 academic year averaged normalized gains of 43% and 52%, respectively. These results compare very favorably with Hake's findings for the most successful innovative classes around the country. The spring semester class also posted normalized gains on the FMCE that were nearly four times higher than their peers in the regular classes. This is not just an instructor effect as can be seen from the FCI results plotted in Figure 1. The same instructor taught all four courses at NCSU in chronological order. Even though the instructor was aware of active learning techniques and the benefits of active learning, when he taught the mechanics class in a traditional lecture format the FCI normalized gains dropped to a value comparable to the national average for traditional lecture courses.

The second semester course concept test results indicate that there is some instructor effect. The normalized gains from the SCALE-UP classes on DIRECT and CSEM in the 1998 fall semester are only slightly better than from the regular classes at NCSU. These results suggest that student learning gains were not significantly better than the regular lecture classes. However, the second semester SCALE-UP courses were taught with more lecture and fewer activities than the first semester courses. This past semester, the second semester SCALE-UP class was taught by a different instructor who used more group activities with less lecturing. The CSEM normalized gains were more than twice that of the regular classes and there was some improvement on the DIRECT test as well.

Table 3: Diagnostic tests results from the SCALE-UP project. The figure of merit *h* is the normalized or fractional pre/post gain (see description in the text of *h* in section II). Note that for each of the four concept tests, the SCALE-UP classes at North Carolina State University are achieving significantly greater normalized gains that the traditional lecture classes, in most cases the normalized gains are least 2 times larger for the SCALE-UP students.

	h from 1st Semester classes: Mechanics	
	FCI	FMCE
Traditional Lecture classes at NCSU	0.21	0.11
F98 SCALE-UP	0.42	
S99 SCALE-UP	0.52*	0.39

	h from 2nd Semester classes: Electricity, Magnetism, & Optics	
	CSEM	DIRECT
Traditional Lecture classes at NCSU	0.14	0.10
F98 SCALE-UP	0.21	0.17
F99 SCALE-UP	0.36	0.21*

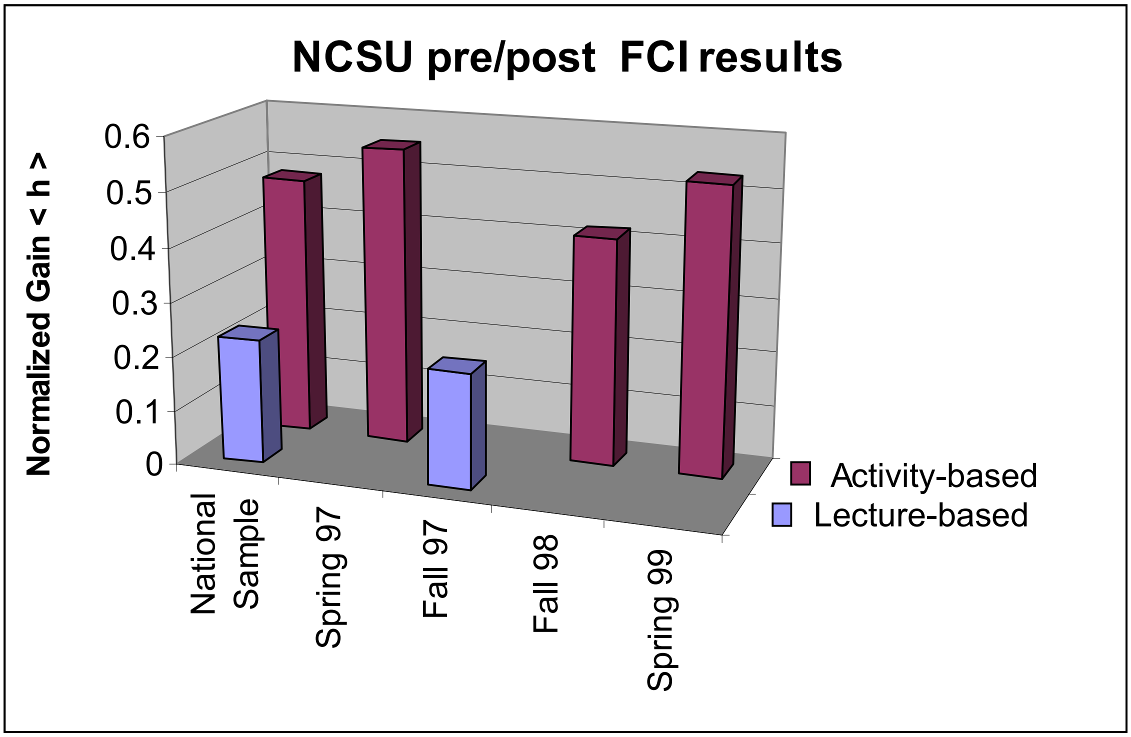
* indicates *h* was calculated using an estimated pre-course average because only post-course data was collected. Pre-course value estimated conservatively using correlation with other diagnostic tests and pre-test values from previous semesters.

Exam Problems

During the 1998-99 academic year, the first semester SCALE-UP students outperformed the students in the regular lecture sections of the class on exam problems written for the regular class. In the fall semester, the SCALE-UP students outperformed their peers 88% of the time. During the spring semester, the SCALE-UP students had higher scores 69% of the time. In general, when the traditional students did better, the problems tended to be one-step problems like simple unit conversions and the performance differences were smaller.

In addition, the students in both the first and second semester SCALE-UP classes performed well on qualitative and complex quantitative problems. The students learned to reason qualitatively and to write short essays using physics concepts without calculations. In general, they demonstrated a high level of understanding. The SCALE-UP students' performance on all three types of exam problems suggests a better understanding of the main concepts.

Figure 1: FCI results from one instructor teaching lecture and SCALE-UP activity-based introductory physics classes at North Carolina State University. Note that the results for both lecture and activity-based classes are comparable to the national sample from Hake's 6000 student study in R.R. Hake, "Active engagement vs. traditional methods: A six thousand student study of mechanics test data for introductory physics courses," *Am. J. Phys*. **66** (1), 64-74 (1998).



B. Objective: Developing Expert-Like Problem Solving Skills

In the previous section we saw that in general, the SCALE-UP students were more successful problem solvers than the regular lecture students. To help students become at better solving problems more complex than typical textbook exercises, it is necessary to help students move from a means-end formula-based novice approach to a more expert-like problem-solving approach.

To achieve this goal, both classes were instructed in the use of an expert-like heuristic problem solving strategy called GOAL after the four basic steps: Gather, Organize, Analyze, Learn.[10] A description of the GOAL problem solving strategy and some examples of typical complex or "Real World" quantitative problems used in the SCALE-UP classes can be found in the invited paper, "Promoting collaborative group in large enrollment classes," in session 1526. Using this expert-like problem solving strategy allowed students to successfully solve these Real World problems, which are normally considered too difficult for introductory physics classes. Most students were successful with GOAL problems on tests and quizzes. The mean scores for the GOAL problems were often higher than for easier, more traditional problems. This indicates that a majority of the students were able to attempt and make progress on problems normally considered too difficult for students in the regular classes.

A small number of students in the first semester SCALE-UP and lecture classes participated in problem solving interviews. From a preliminary analysis of the interviews, all but one of the regular lecture students used a “plug-and-chug”, “find-the-formula-in-the-book” approach. The students had difficulty evaluating their answers and if their approach did not work, the students were at a loss for what to do next. They displayed typical characteristics of novice problem solvers.[11] Although the SCALE-UP students did not formally use the GOAL strategy to solve problems, they did display some characteristics of expert-problem solving behavior. These characteristics included identifying the main physics principle at work in the problem, coming up with a plan (although vague and often incomplete) to approach the problem, and checking to see if their answers were reasonable.

C. Objective: Improving Communication, Teamwork, and Questioning Skills

We use classroom observations, student presentations, and student evaluations of how well their group performed on a particular activity to see if students were improving their communication, teamwork, and questioning skills. This objective is probably the hardest to evaluate objectively. To validate our findings, we compared them with those of outside observers and evaluators. By midsemester, classroom observers and outside evaluators commented that during group activities the students were largely on task, engaged, and working well in their groups. In addition, the observers and evaluators also noted that during class discussions, the SCALE students were asking significantly more thoughtful questions about the topics under discussion than students in the regular classes.

A number of steps have been taken to help the students form successful groups including having the groups write a contract of what is expected of all group members and bonus points on exams if the group average is 80% or better. A description of the steps taken to promote successful groups, including examples of group contracts, can be found in the invited paper, “Promoting collaborative group in large enrollment classes,” in session 1526. In the 1999 fall semester, the SCALE-UP students switched groups in the middle of the semester for the first time. In the focus group interviews, students commented that while the change was traumatic at first, they got used to their new groups after three or four weeks. Not surprisingly, the second set of group contracts showed some improvement. More groups mentioned making sure that all members of the group understand the material and respected the other members of the group.

D. Objective: Developing Laboratory Skills

In regular physics classes, while most students learn to carry out the experiments given by the cookbook-style lab procedure and prepare a formal lab report, very few students develop the more advanced skills needed to design, carry out, analyze, and evaluate an experiment, one of the desired ABET outcomes for program graduates. In fact, few students have a good grasp of what it means to make a measurement.[12] In SCALE-UP, laboratory activities are being developed to help students develop these skills.

For the SCALE-UP classes, student laboratory skills are evaluated by their performance on the following activities:

1. lab practica – quizzes where students are asked to demonstrate the skills need for previous lab activities.
2. open ended-lab activities where students are asked a question and must design, carry out, and analyze an experiment to answer it without being given a specific procedure to follow.

The lab practica are useful for determining if each student has mastered basic measurement and analysis skills required for completing previous lab activities. The practica provide motivation for each student to know how to do each part of the lab activities. Students have been very successful at demonstrating mastery of basic skills; in the next semester we will be adding questions on the main physics concepts underlying the lab to the practica to see if students are making this connection.

Student performance on the open lab activities has been mixed; only a few groups have been able to complete these labs without substantial guidance from the instructors. Based on these results, we are redesigning the lab activities in both classes to help students learn more about measurement and designing experiments in stages over the semester.

In addition, we have instituted peer evaluation of lab reports during the 1999 fall semester. Students are asked to rank five conclusions written by other student groups or the TAs. The students are graded based on how well their rankings match the TAs. This type of evaluation is a higher order cognitive skill. We are currently studying the conclusions from this semester and comparing them with previous semesters to see if this activity helps students write better conclusions.

E. Objective: Learning to use technology

Laptop computers and the web are used extensively in both SCALE-UP courses. Each student group of three or four students uses a laptop with access to the Internet. Students have been observed to use computers to hunt for information on the web, to collect data from experiments and to run simulations, to display and submit homework, and to create mathematical models of physical situations using spreadsheets. In addition, group lab reports are prepared using word processors. There is a tendency for the student groups to allow the person with the most computer expertise to operate the computer. We find that by making sure students trade-off who is using the computer and making sure everyone can pass the lab practica, even students initially uncomfortable with computers get to learn how to perform the operations described above.

F. Objective: Positive Physics Learning Experience

Many students dislike introductory physics courses, see physics as very formula oriented and boring, or find physics extremely difficult. To evaluate students' physics learning experience, we looked at whether a higher fraction of students passed the course and whether they thought they benefited from or preferred the SCALE-UP approach more than the traditional lecture approach. Overall the SCALE-UP students had a lower failure rate than traditional lecture students (13% vs 25%, respectively), and in particular, the failure rates for women and minorities[13] (women 9%. vs 27%, minorities 8% vs 48%) were much less than those found in the regular sections at NCSU.

Course evaluations and focus group interviews were used to learn more about student perceptions of the class. The results have been quite encouraging. The SCALE-UP students' comments can be summarized in the following points:.

- Students felt that SCALE-UP required more work than a regular class but believed the extra effort was worth it
- They liked in-class group work, but not out-of-class group work
- They liked that instructors knew their name & progress
- SCALE-UP is more effective for learning than regular lecture class
- The SCALE-UP class doesn't put them to sleep; it is harder to remain invisible
- The students recognized that the SCALE-UP classes emphasized understanding more than the regular lecture classes.

The positive nature of students' experience with SCALE-UP is illustrated in the following quotes from the focus group transcripts.

> " I can deal with the lecture class, it is just that I enjoy more ... getting more into the interactive projects. It is more hands on. If you don't understand something you just ask the guy next to you, nobody yells at you for talking."
>
> " I actually know how I learn through the SCALE-UP physics…through the way that it is set up, through the way they taught us by solving problems. It helped me to learn not so much to get an answer but to actually understand concepts. I also apply that to the rest of my classes. I think from now on, I will do a lot better in my classes just by taking this class–through all the teaching we learned how to solve problems and think through problems."
>
> "I have studied for a test with some of the traditional 205 students and like they always point to the book for everything, like looking for a formula for everything. [The instructor] makes sure that we understand the concept, we can almost derive the formula for whatever we need. And we seem to understand more of the aspects of physics. I definitely feel that, compared to traditional, we have a more in depth understanding and knowledge of what is going on."

IV. Conclusion

In this paper, we have shown how course evaluation tools can be used to determine if a course is making progress in achieving ABET-like learning objectives. The Phase II SCALE-UP introductory physics classes are clearly doing a better job of meeting our pre-defined learning objectives than the NCSU regular lecture classes of the same course. Evidence of this success includes:

- Normalized gains of SCALE-UP classes were larger, often 2-4 times larger, than those of lecture classes on standardized concept tests,

- SCALE-UP students outperformed their peers on exam problems written for the regular classes,
- SCALE-UP students demonstrated signs of expert problem-solving behavior while solving problems in class and in interviews considered too complex for students in regular classes at NCSU,
- The SCALE-UP students were observed to work well in groups on activities and to ask more thoughtful questions than students in regular classes, and
- While the SCALE-UP students think their class is more work than a regular section, they believe it is worth it because it helps them learn physics better.

The evaluation tools used in this paper are good enough to provide feedback to point out areas that need further improvement. ABET refers to this process as closing the loop in iterative curriculum development, an important step in the evaluation process. Our evaluations suggest three areas in particular where we might focus our efforts to continue to improve the SCALE-UP curriculum: conceptual understanding, problem solving, and laboratory skills. First, while the performance of the SCALE-UP students on concept tests show significant improvement, the gains indicate that SCALE-UP students are only achieving 20-50% of what is possible. We are doing an item analysis on the diagnostic tests and on exam problems to identify particular concepts that are still giving students difficulty and modifying our curriculum to address them better. Second, student problem-solving has been improved but could use more work. Many students are not spontaneously using our GOAL problem solving approach to solve an unfamiliar difficult problem. This suggests that SCALE-UP instructors need to do more modeling and coaching of the GOAL approach in the first half of the semester. Finally, student performance on open-ended lab questions indicate that the SCALE-UP curriculum could do better in terms of helping students develop the skills to design, carry out, analyze, and evaluate an experiment when they are not given a written procedure. We are in the process of modifying the laboratory activities to help students develop these skills over the course of a semester.

SCALE-UP is an ongoing project at NCSU. We are looking at helping interested institutions adapt it for their own use. If you are interested, please visit our website at www.ncsu.edu/PER or contact Robert Beichner by e-mail at Beichner@ncsu.edu. SCALE-UP instruction materials are available from a password protected portion of the website.

V. Acknowledgments

We would like to thank the U.S. Department of Education FIPSE program (PB116B71905), Hewlett Packard, and the National Science Foundation (DUE-9752313) for their support. We would also like to thank Melissa Dancy, Scott Bonham, Margaret Gjertsen, and John Risley for their contributions to the SCALE-UP project.

Bibliography

1. L.C. McDermott, "Research on conceptual understanding in mechanics," *Phys. Today* **37** (7), 24 (July 1984); I.A. Halloun and D. Hestenes, "The initial knowledge state of students," *Am. J. Phys*. **53** (11), 1043-1055 (1985); I.A. Halloun and D. Hestenes, "Common sense concepts about motion," *Am. J. Phys*. **53** (11), 1056-1065 (1985); D. Hammer, "Two approaches to learning physics," *Phys. Teach*. **27** (9), 664-670 (1989); A. van Heuvelen, "Learning to think like a physicist: A review of research based instructional strategies," *Am. J. Phys*. **59**, 898-907 (1991); A. Van Heuvelen, "Overview, case study physics," *Am J. Phys*. **59** (10), 898-906 (1991); L.C. McDermott, "Millikan Lecture 1990: What we teach and what is learned — Closing the gap," *Am. J. Phys*. **59** (4), 301-315 (1991); D. Hammer, "Epistemological beliefs in introductory physics," *Cognition and Instruction* **12**, 151-183 (1994); D. Hammer, "Students' beliefs about conceptual knowledge in introductory physics," *Int. J Sci. Ed.* **16**, 385-403 (1994); E. Mazur, *Peer Instruction: A Users Manual* (Prentice Hall, New Jersey, 1997), 3-18.

2. Other examples include L.C. McDermott and P.S. Shaffer, *Tutorials in Introductory Physics* (Preliminary edition) (Prentice Hall, Upper Saddle River NY, 1997); P.W. Laws, *Workshop Physics Activity Guide* (Wiley, New York NY, 1997); and University of Minnesota's *Cooperative Group Problem Solving* curriculum (information on this curriculum is available online at http://www.physics.umn.edu/groups/physed/). Evaluations of these curricula can be found in the following references: (***Tutorials in Introductory Physics***) G.E. Francis, "Effectiveness of Tutorials in Introductory Physics," in *AIP Conference Proceeding No. 399 The Changing Role of Physics Departments in Modern Universities: Proceedings of the International Conference on Undergraduate Physics Education*, edited by E.F. Redish and J.S. Rigden (AIP Press, Woodbury NY, 1997), 567-574; E.F. Redish, J.M. Saul, and R.N. Steinberg, "On the effectiveness of active-engagement microcomputer-based laboratories," *Am. J. Phys.* **65** (1), 45-54 (1997); K. Wosilait, P.R.L. Heron, P.S. Shaffer, and L.C. McDermott, "Addressing student difficulties in applying a wave model to the interference and diffraction of light," *PER Supplement to Am. J. Phys.* **67** (7), S5-S15 (1999); (***Workshop Physics***) P. Laws, "Calculus-based physics without lectures," *Phys. Today* **44 (**12), 24-31 (December, 1991); P. Laws, "Millikan Lecture 1996: Promoting active learning based on physics education research in introductory physics classes," *Am. J. Phys.* **65** (1), 13-21 (1997); (***Cooperative Group Problem Solving***) P. Heller, R. Keith, and S. Anderson, "Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving," *Am. J. Phys*. **60** (7), 627-636 (1992); P. Heller and M. Hollabaugh, "Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups," *Am. J. Phys*. **60** (7), 637-644 (1992).

3. E.F. Redish, "Implications of Cognitive Studies for Teaching Physics," *Am. J. Phys.* **62** (9), 796-803 (1994).

4. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," *Phys. Teach.* **30** (3), 141-158 (1992).

5. Over six thousand students participated in a study of FCI results by Hake. See R.R. Hake, "Active engagement vs. traditional methods: A six thousand student study of mechanics test data for introductory physics courses," *Am. J. Phys*. **66** (1), 64-74 (1998).

6. Hake found that unlike average post score, the average normalized gain h did not correlate with the average pre-test average. This makes h a more useful measure for comparing classes with significantly different pre-course scores. See Ref. 5.

7. J.M. Saul, *Beyond Problem Solving: Evaluating Introductory Physics Courses Through the Hidden Curriculum*, Ph.D. Dissertation, University of Maryland, College Park, 1998 (unpublished).

8. In addition, the last two weeks of the course are used as a very brief introduction to modern physics.

9. DIRECT was developed, validated, and tested for reliability by graduate student in the PER-D group as a Ph.D. project. For more information, see P.V. Engelhardt, *Examining Students' Understanding of Direct Current Electrical Circuits Through Multiple- Choice Testing and Interviews*, Ph.D. dissertation, North Carolina State University, 1997 (unpublished).

10. A description of the GOAL approach can be found in R.A. Serway and R.J. Beichner, *Physics for Scientists an Engineers*, 5th ed. (Saunders College Publishing, Orlando Fl, 2000), p. 47. GOAL in a mnemonic for the often-used four step problem solving framework of understanding the problem, devising a plan, carrying out the plan, and looking back espoused in G. Polya, *How to Solve It* (Doubleday, Garden City NY, 1945).

11. D.P. Maloney, "Research of problem solving: physics," in *Handbook of Research on Science Teaching and Learning*, edited by D.L. Gabel (Macmillan Publishing Company, New York, 1994), 327-354.

12. D. Deardorff and R.J. Beichner, "How much do students really learn from physics labs," Contributed paper presented at the fall meeting of the North Carolina Section of AAPT, Boone NC (1999).

13. In this case, the term "minorities" is used to describe African-American, Hispanic, and Native American physics students

JEFFERY M. SAUL
Jeff Saul is a post-doctoral research assistant in the physics education research and development group at North Carolina Sate University. He received a B.S. degree in Applied Physics from University of California at Irvine and received the first Ph.D. in Physics in physics education research from University of Maryland under Edward F. Redish. His current research interests are in student learning of physics and development of active-learning introductory physics curricula. In addition, he is an author of the instructor's manual to *Physics for Scientists and Engineers* by Ray Serway and Robert Beichner.

DUANE L. DEARDORFF
Duane Deardorff is currently a graduate student at North Carolina State University pursuing a Ph.D. in physics. He received a B.A. in physics with distinction from Manchester College. His current research interests are in student learning of physics in introductory physics classes, with an emphasis on studying student understanding of measurement and uncertainty. In addition, he is a contributing author for the student study guide and the instructor's manual to *Physics for Scientists and Engineers* by Ray Serway and Robert Beichner

DAVID S. ABBOTT
David Abbott is currently a graduate student at North Carolina State University pursuing a Ph.D. in physics. He received a B.S. degree in Physics from University of Delaware. After receiving his M.S. degree from University of Virginia, David taught physics as adjunct faculty at Delaware Technical Community College and other community colleges for five years. He is pursuing his Ph.D. dissertation research under a National Science Foundation grant to study use of technology for improving student learning in introductory physics classes.

RHETT J. ALLAIN
Rhett Allain is currently a graduate student at North Carolina State University pursuing a Ph.D. in Physics. After receiving his M.S. degree in Physics from the University of Alabama, Rhett was an assistant professor at Judson College in Marion, Alabama. His current research interests are in student learning of physics in introductory physics courses, particularly student understanding of electric potential difference and developing student activities using computer-assisted peer evaluation. In addition, he is a contributing author for the instructor's manual to *Physics for Scientists and Engineers* by Ray Serway and Robert Beichner.

ROBERT J. BEICHNER
Robert Beichner is Associate Professor of Physics at the North Carolina State University. He received B.S. degrees in Physics and Mathematics from Pennsylvania State University and a M.S. in Physics from University of Illinois. His Ph.D. from SUNY at Buffalo is in Science Education. His research interests are in student learning of physics. He recently completed work on the 5th edition of Serway and Beichner, *Physics for Scientists and Engineers*, a leading introductory textbook.