The SCALE-UP Project:
A Student-Centered Active Learning Environment for Undergraduate Programs

Robert J. Beichner, North Carolina State University, Raleigh, NC

I. Introduction

SCALE-UP[1] stands for "Student-Centered Active Learning Environment for Undergraduate Programs." It describes a place where student teams are given interesting things to investigate while their instructor roams—asking questions, sending one team to help another, or asking why someone else got a different answer. Even in a science course, there is usually no need for a separate lab. Most of the "lectures" are class-wide discussions. The groups are carefully structured and give students many opportunities to interact with each other and the instructor. Three teams (labeled A, B, and C) sit at each round table and have white boards nearby. Every group has a laptop for searching the web. At NC State, the original site, classes usually have 11 tables of nine students. Most of the 50+ schools that have adopted the approach have smaller classes, while a few have even larger ones[2]. The majority of class time is spent on 10 or 15 minute "tangibles" and "ponderables." Essentially these are hands-on activities, simulations, or interesting questions and problems. In science classes there are usually some longer, hypothesis-driven lab activities where students have to write detailed reports. Occasionally there will be lecturing, but that is mostly to provide motivation and a view of the "big picture," which can be difficult for students to discern when they are not familiar with the entire course content.

Figure 1. This photograph from Ithaca College shows the typical SCALE-UP environment: round tables seating three teams of three students, surrounded by screens, white boards and handy equipment storage. There is a teacher station located somewhere in the midst of the action. (Note that these tables are slightly smaller than usual.) Photo courtesy of Michael Rogers.

Social interactions between students and with their teachers appear to be the "active ingredient" that makes the approach work. As more and more instruction is handled virtually via the web, taking advantage of the relationship-building capability of the real people in brick and mortar universities becomes even more important. The most quoted study[3] in all of higher education research indicates that we probably have it right: "What Matters in College" are the relationships students build with each other and with their teachers. The research base that supports the design of SCALE-UP has been culled from many sources. The fundamental approach of active, collaborative, social learning has been reported in hundreds of studies[4]. Studio-based learning is not new, but its application to science classes is fairly recent[5].

Physics, chemistry, math, biology, astronomy, engineering, and even literature courses have utilized this approach. A political science class is in development at one adopting school. The teacher could pick some current event to focus the students' attention, for example a government official's Congressional testimony on some controversial topic. The "A" group at each table would go to the web to see how CNN covered the event. The "B" groups would read the Washington Post coverage, while the "C" groups could find the Fox News website. Then they would compare and see what aspects were covered by all three and which things were missing in some. They might then be sent on a search to find the least biased presentation, perhaps by an international organization like the BBC. Whether the topic is current events or chemistry, the basic idea is the same[6]. Student teams work on interesting tasks while teachers coach.

Some people think the rooms look more like restaurants than classrooms. Like an eating establishment, the spaces are carefully designed to facilitate interactions between people. They are

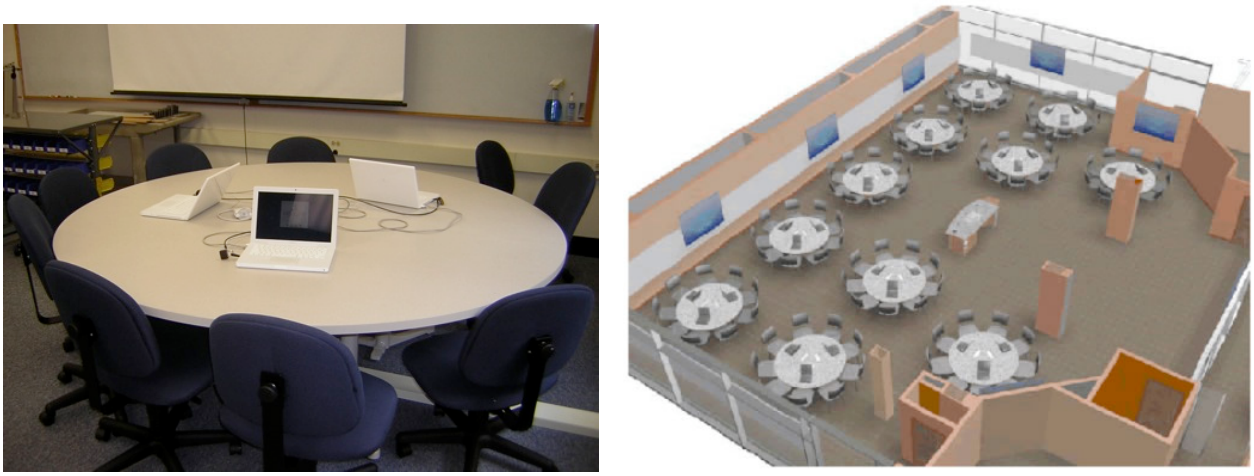

Figure 2. Classrooms at Pitt and MIT illustrate the common features of most SCALE-UP classrooms: 7' diameter round tables, whiteboards, projection screens, and one laptop/team. Photo courtesy of Adam Leibovich. Graphic courtesy of John Belcher.

definitely noisy places, with lively conversations going on nearly all the time. For larger classes, a teaching assistant provides additional help. The instructor typically wears a wireless microphone to make it easier to gain everyone's attention for classwide discussions. Often students working on an activity will skip their break in the middle of a two-hour class so they can continue "pondering" an intriguing question. A decade of research indicates significant improvements in learning.

II. Evidence of Efficacy

Rigorous evaluations of learning have been conducted, either in parallel with curriculum development and classroom design work, or as a follow-up to such efforts. Many adopters have given conceptual learning assessments (using nationally-recognized instruments in a pretest/posttest protocol), and collected portfolios of student work. Several schools have conducted student interviews and collected information from focus groups, supplementing hundreds of hours of classroom video and audio recordings made at NC State during the early development phases. More details of the research behind the room design as well as outcomes of studies of educational impact are available[7].

Concept learning

There is ample evidence[7] from multiple adopting sites that students in SCALE-UP classes gain a better conceptual understanding than their peers in traditional lecture-based classes. As Figure 3 shows for the first and second semesters of introductory physics, students performed better on a variety of conceptual surveys. The pattern apparent in Fig. 3(b), where students in the top third of their class made the most progress toward perfect scores on the assessment tests, is an important counter-argument to those who complain that "reform courses only benefit the weaker students and we are ignoring the stars of tomorrow." Clearly that is not the case.

Other schools have had similar results. At Florida State, normalized gains on the FCI from the first (Spring 2008) and second (Summer 2008) implementations of General Physics were approximately 50%[8], far surpassing the typically seen 23% for traditional courses[9]. Florida International University notes[10], "These courses have been extremely successful, in terms of student learning outcomes, faculty assessments, and recruiting. The average student performance on the Force Concept Inventory (FCI) in the modeling-based [studio physics] courses is roughly a factor of 2.5 better than in our traditional courses." At Penn State-Erie, over 550 students have enrolled in SCALE-UP physics, as of the summer of 2008. Scores on the FCI post-test have increased from an average score of 46%

correct before SCALE-UP to 74% correct since SCALE-UP began[11]. The University of Pittsburgh reports[12] what they call "striking" gains on a test[13] of electricity and magnetism concepts. Positive impacts are manifested in other areas as well. Chemistry faculty have published[14] findings of learning gains. An internal report[15] on the Engineering Statics course at Clemson reports, "One of the common concerns expressed by my colleagues is that I must not be covering as much material since I am using class time to complete activities. My response is that I cover the same amount of material as other instructors." NC State notes[7] the same situation in Physics. Biology learning is being studied at the University of Minnesota, Florida Gulf Coast University, and the University of Colorado. Minnesota Biology professor Robin Wright has been so successful with her classes that she believes[16] the SCALE-UP approach would work with up to 250 students at a time. Elizabeth Wolfe completed her University of Victoria MS thesis[17] doing a study of learning in a SCALE-UP computer database systems course and reports that the "evaluation surpassed expectations both with regard to course delivery and student perception of teamwork." The NC State study[7] also examined teacher effects at two schools and found that students of teachers in SCALE-UP settings had greater conceptual learning gains than students of those same teachers in lecture settings. MIT has carried out several studies and reports improved[18] conceptual learning and significantly better long-term retention[19] of those gains.

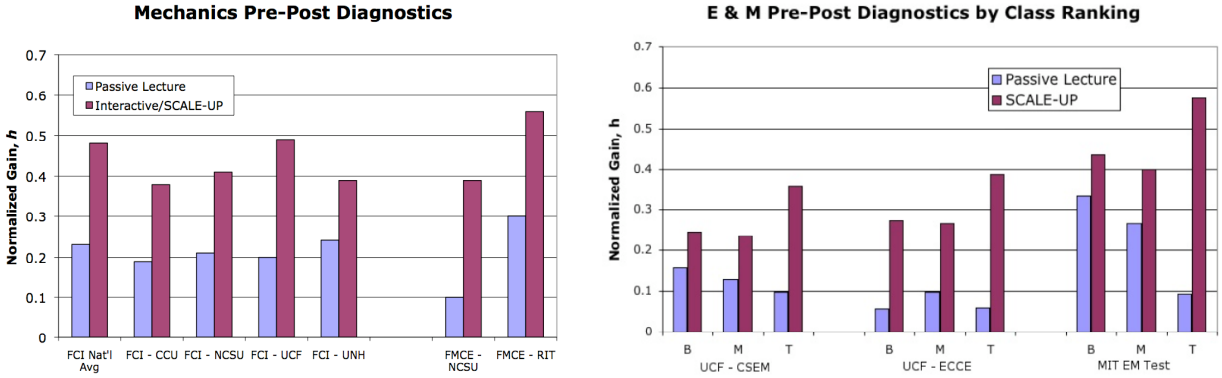

Figure 3: (a) SCALE-UP students demonstrated better improvement in conceptual understanding than Lecture/Lab classes by achieving higher normalized gains for the Mechanics semester pre/post force and motion concept tests at Coastal Carolina University (CCU), North Carolina State University (NCSU), University of Central Florida (UCF), University of New Hampshire (UNH), and Rochester Institute of Technology (RIT). FCI is the Force Concept Inventory developed by Hestenes, *et al.*[20] FMCE is the Force and Motion Conceptual Evaluation developed by Thornton and Sokoloff.[21] The FCI national average is from Hake's 6,000 student study comparing Interactive Engagement classes with traditional Lecture/Laboratory classes.[9] (b) B, M. and T stand for Bottom, Middle, and Top thirds of the class, as measured by conceptual pretest scores. Students in the top third of their classes had the highest normalized gains, possibly because they were teaching their peers. CSEM is the Conceptual Survey of Electricity & Magnetism developed by Maloney, *et. al.*[22] ECCE is the Electric Circuit Conceptual Evaluation[23] developed by Thornton and Sokoloff. The MIT E & M test was developed at MIT for their SCALE-UP implementation.[18]

Skill Development

Since nearly all SCALE-UP courses across the country have been in the STEM (Science, Technology, Engineering, and Math) fields, most schools have been very interested in the impact of the approach on measurement and problem solving skills, as well as communication and teamsmanship. These have been evaluated at several places. Work at NC State showed that SCALE-UP students' lab measurement skills improved[24] and they achieved one letter grade better on tests written by lecturers than did the lecturers' own students[7].

Streitmatter[25] reported that female students prefer, and achieve better in, classrooms where learning activities are structured as cooperative endeavors rather than within a competitive structure. It is interesting to note the progress of women students at Penn State-Erie, where they "have SAT math scores and mathematics placement test scores that are well below those of their male counterparts ($p < .001$ and $p < .05$ respectively). At the time of the first test of the semester, females still have significantly different scores, with an average of 62% versus the male mean of 74% ($p < .001$). By the second exam though, the females catch up to the males, and maintain this equality of achievement through the final examination ($p > .05$). The final course grades of males and females in SCALE-UP are not significantly different ($p > .05$), despite the fact that women start the course with lower scores on tests of prerequisite skills[26]."

Affective outcomes

At schools where they have a choice, students almost always prefer SCALE-UP based classes compared to lecture courses. Although attitudes of SCALE-UP students are sometimes studied directly[14,18], they are more often revealed indirectly; for example, students universally select the SCALE-UP version for their second semester course, they report their friends direct them into SCALE-UP classes, SCALE-UP sections fill before lecture sections, etc. NC State has a five-year average attendance rate of more than 90%, even though attendance is not required for SCALE-UP classes[7]. At the very least, this implies that students value classtime. At best it may also indicate they enjoy learning in this type of setting. Minnesota's evaluation[27] of their pilot classrooms notes, "The instructors who were interviewed enjoyed teaching in the rooms so much that their only concern was a fear of not being able to continue to teach in these new learning spaces. Similarly, more than 85 percent of students overwhelmingly recommended the Active Learning Classrooms for other classes."

Retention rates are a sort of "grand total" of educational outcomes, not the least of which is student motivation. For example, the DFW rate (drop, fail, withdraw) for studio-based modeling classes at

Florida International University is "1/4th the DFW rate in traditional classes. Faculty evaluations and student feedback have been overwhelmingly positive, and the courses are drawing roughly four times the room capacity in requests to enter the class. We also find 10-20% of the students pursue physics minors and majors after taking the course, either adding a second major/minor or switching majors[10]." NC State found similar results from data[7] comparing pass/fail rates for nearly 16,000 traditional and SCALE-UP students taking physics. Failure rates, especially for women and minorities, are reduced by a factor of four or five, as seen in Figure 4. This is consistent with Colbeck's findings[28] that women's confidence increases when clear expectations are presented and men's confidence increases with increased faculty interaction. She also found high correlation between collaborative learning and student confidence for both genders. Bandura's theory[29] of social cognition implies that this increased confidence will lead to improved performance and more resilience in a challenging environment.

At Clemson, "beginning in Fall 2006, all freshman Calculus I courses were taught using the SCALE-UP model, in order to address high DFW rates. Historically, the DFW percentage was 44%, and had seen a sharp increase prior to Fall 2006 in most freshman calculus classes. The current DFW rate for all these courses, which includes nearly 800 freshmen, has dropped to approximately 22% in that program, which is encouraging our faculty to adopt the SCALE-UP approach permanently as part of our academic culture.[30]"

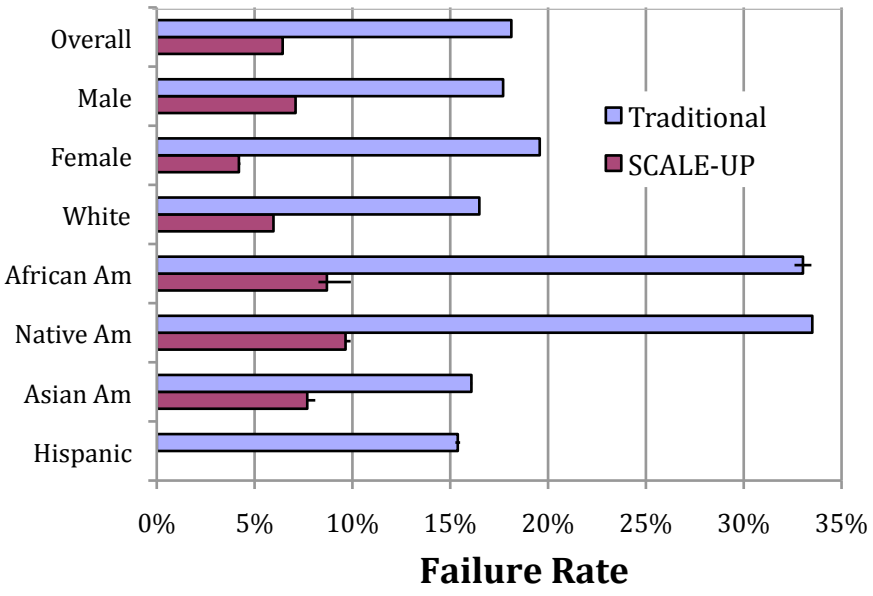

Figure 4: Failure rate comparison for NCSU Physics I & II classes. Here, failing means receiving a grade lower than C – in the mechanics course or less than a D – in the E & M course, the grades needed to receive credit for taking the course. No Hispanic SCALE-UP students failed during the five years of data collection from 16,000+ students, so that bar is has zero length. Error bars represent standard error of the mean, and are mostly too small to be seen.

A direct way to measure student attitudes about a course is to simply ask them. Two visitors to the NC State project interviewed a focus group of students who had completed a lecture course for their first semester of physics and a SCALE-UP section for the second physics class. In their report, they noted,

> *They* [the students] *felt they were learning the material at a deeper conceptual level in Scale-Up as compared to the lecture format, and that there was much less rote memorization on their part. They felt that the contributing factors to this positive outcome were the hands-on nature of the classroom experience, the collaborative work format ("I learn much better from my peers than from my Professors"), and the availability of faculty and TA's for interaction during class...*

Other outcomes

Several schools have looked at student performance in later courses. At NC State, there has been no change[7] in the overall DFW rate for engineering statics courses, even though the SCALE-UP physics failure rate is approximately 1/3 what it was with traditional classes. In fact, students defined as "at risk" (based on SAT math scores < 500) fail statics courses 17% of the time if they took a SCALE-UP physics course, but 31% of the time if their background included only lecture-based physics courses.

Students seem to recognize the value of what they are learning. When Penn State-Erie conducted[31] an opinion survey of former students, they found that the problem-solving, communication, and teamwork skills learned in SCALE-UP classes were being utilized in other courses. A computer & software engineering student noted, "Teamwork and group problem-solving…are very important skills for engineers and other fields requiring group collaboration." A mechanical engineering student stated, "SCALE-UP physics helped me to learn that exploring concepts on my own, outside of a lecture, helps me to remember them better."

Faculty, too, are seeing that changing their focus from teaching to learning is important. As noted by Peter Dourmashkin, who teaches physics at MIT, "Traditionally, in large lectures, you do what is possible to do in front of 500 people, not because it's what you should do. Now we're asking the question: What do we really want our students to learn about electricity and magnetism?"

III. Assessment methods

When starting on a journey, it is always good to know where you want to go. Similarly, when taking on the wholesale redesign of college-level instruction, it is important to have clearly delineated objectives. Obviously these will depend on the type of course being reformed.

Before the pilot classroom at NC State was built, the project team sat down and outlined their large-scale objectives for the two-semester introductory physics sequence, with the aim of tying them

tightly to the accreditation criteria established recently by ABET[32], the Accreditation Board for Engineering and Technology. These objectives are reproduced in Table 1. It is important to realize that each of these overarching objectives had numerous sub-objectives that were behavioral in nature. In other words, they explicitly described something the student should be able to do. For

NCSU SCALE-UP Objectives for Calculus-based Intro Physics	ABET 2000 Criterion 3: Program Outcome Requirements	Assessment Method
SCALE-UP students should:	Engineering programs must demonstrate their students have:	
develop a good functional **understanding** of physics.	(3a) an ability to apply knowledge of mathematics, science, and engineering	Pre/Post concept tests on forces, graphs, electromagnetism, circuits
begin developing expert-like **problem solving** skills.	(3e) an ability to identify, formulate, and solve engineering problems	Student portfolios; interviews; test comparisons to control groups
develop **laboratory** skills.	(3b) an ability to design and conduct experiments, as well as to analyze and interpret data	Practical testing; student portfolios
develop **technology** skills..	(3k) an ability to use the techniques, skills, and modern tools necessary for engineering practice	In-class observations via field notes; practical testing; student portfolios
improve their **communication**, interpersonal, and questioning skills.	(3d) an ability to function on multi-disciplinary teams (3g) an ability to communicate effectively	In-class observations via field notes, audio, and video recording; interviews; focus groups
develop **attitudes** that are favorable for learning physics.	(3h) the broad education necessary to understand the impact of engineering solutions in a global and societal context (3i) a recognition of the need for, and an ability to engage in life-long learning	Maryland Physics Expectations (MPEX) survey; interviews; in-class observations via field notes
have a **positive learning experience**.	Not in the ABET criteria	Course evaluations; interviews; focus groups

Table 1: Overall objectives for NCSU SCALE-UP Physics and their relationship to the ABET criteria, which are indicated by (3a), (3b), etc. in the middle column. Pre-post concept tests were the FCI[20], FMCE[21], TUG-K[33], CSEM[22], BEMA[13], and DIRECT[34]. MPEX stands for Maryland Physics Expectation Survey[35].

example, the first objective looked for understanding of physics. This could be demonstrated by a student who could:

A. describe and explain physics concepts including knowing where and when they apply

B. apply physics concepts when solving problems and examining physical phenomena

C. apply concepts in new contexts (transfer)

D. translate between multiple-representations of the same concept (for example: between words, equations, graphs, and diagrams)

E. combine concepts when analyzing a situation

F. evaluate explanations of physical phenomena

Each of these sub-objectives then had even finer grained objectives that explicitly invoked the physics content. This made it easier to utilize portfolios of student work to assess whether the objective had been met.

The Reformed Teacher Observation Protocol (RTOP) has also been used to evaluate SCALE-UP classrooms. As MacIssac and Falconer state[36], "The RTOP instrument is designed to constructively critique details of classroom practices including cooperative learning, interactive engagement, and

certain types of PER [Physics Education Research] activities, as well as findings collectively known as pedagogical content knowledge." It has been used in more than 400 science and mathematics classrooms and found to correlate strongly with student learning gains (r = 0.70-0.95). Traditional university lecture scores are 20 (out of 100) while a medium-sized lecture using Peer Instruction[37] with an electronic student response system scores from 65-75. Saul[38] used the protocol on two lecture classes and one SCALE-UP class at the University of Central Florida and found the lecture classes scored 35 while the SCALE-UP class received a score of 80.

At NC State, we hired an external evaluator to compare our traditional lecture-based and SCALE-UP classes. About a lecture class taught by a veteran professor she writes,

> *Students did ask questions in this class; however, most questions were about details of course procedures or questions requesting factual information related to physics. While this instructor attempted to make the large lecture section more interactive, the students did not exhibit behaviors that signaled positive intellectual regard for the activity.*

However, after observing one of the SCALE-UP classes, the external evaluator reported,

> *...students asked over* ***eight*** [emphasis in original] *questions during the period and all of these questions were* ***substantive*** *questions, questions at one of the higher cognitive levels. In addition, students offered explanations about physical phenomena during discussions without being called upon to do so. It is worth remarking that SCALE-UP students began talking to one another about physics even before class began.*

IV. Open questions

One effect we have noticed but have not had an opportunity to explore is the rapid incorporation of members of underrepresented groups into student teams. Although others[39] have found input from minorities and women to be devalued in group settings, that has not been the case in SCALE-UP classes. We would like to know more about the details of interactions taking place within groups. This was examined during the early pilot phase, but not (at least at NC State) since the classes have grown to 99 students.

Others are beginning to study[40] what influences adoption of reforms and how to best approach reluctant faculty. MIT has reported on positive and negative student reactions[18] to implementation of TEAL, their adaptation of SCALE-UP. Pundak and Rozner[41] studied the factors that influenced SCALE-UP adoption at a college in Israel. This is a promising and interesting area that should be pursued.

While studying what influences adoption, it would also be fascinating to track how implementations vary generation by generation. For example, during the early phases of the project, NC State

researchers carefully examined various table shapes, eventually settling on round tables to best facilitate discussions. (In fact, four different diameters were classroom tested.) Several adopters have tried other table geometries. Alabama has X and T-shaped tables. New Hampshire has lollipop-shaped tables, but is now changing to rounds. Tokyo has "bean shaped" tables. It is interesting to note that Tokyo learned about SCALE-UP from the TEAL project at MIT. Georgia Southern, an institution that learned about SCALE-UP from Alabama, has rectangular tables. It would be useful to see why these changes were made and examine their impact on the learning environment.

Figure 5: X and T-shaped tables at the University of Alabama, "Lollipop" tables at the University of New Hampshire, and "Bean" tables at the University of Tokyo. Photos courtesy of Stan Jones, Dawn Merideth, and Toshio Mochizuki.

V. Next steps

I continue to travel to colleges around the country to describe the SCALE-UP learning environment and help others who are considering adopting the approach. Recently many international trips have been made, including visits to schools in Australia, Canada, Chile, China, Italy, Mexico, Portugal, and Switzerland. The SCALE-UP website has been updated so the implementing sites can share their classroom designs, instructional materials, and document the research they have done on the effectiveness of their reforms. Now we need to begin collaborating on educational research projects. It would be useful to see how implementation strategies, classroom designs, course operation, and effectiveness of these adoptions varies by size and type of institution, as well as examine variation due to content or student demographic differences.

Acknowledgements

I would be remiss if I did not thank the hundreds of people who have been involved of developing, adopting, and adapting SCALE-UP for classrooms around the world. In particular, I appreciate the support of my NC State colleagues Phillip Stiles, Jerry Whitten, Dan Solomon, Jo-Ann Cohen, Chris Gould, Michael Paesler, Rich Felder, John Risley, Maria Oliver-Hoyo, Peter Evans, and Everette Allen. Jeff Saul gathered much of the early data on instructional impact and carefully ensured its quality. David Abbott, Rhett Allen, Scott Bonham, Melissa Dancy, Duane Deardorff, and Jeanne Morse also helped with data collection from the initial efforts. For this report, several adopters provided results at a moment's notice or had previously published results which they generously shared, including

John Belcher, Judy Dori, Adam Leibovich, Michael Rogers, Danielle Goodwin, and Aimee Whiteside. External funding was provided by FIPSE (PB116B71905, P116B000659) and NSF (DUE-9752313, DUE-0127050, and DUE-9981107*)*. Of course, none of this would have been possible without the cooperation of all the students taking SCALE-UP classes, especially while we were "working the bugs out." Thanks to one and all.

References

1 The name was originally "Student-Centered Activities for Large Enrollment University Physics." It has been changed because there are now a variety of institutions applying the approach to teach many different content areas to classes of all sizes.

2 MIT and Minnesota have rooms for 117 students. Virginia is planning a room for 162. Minnesota believes they could manage a class of 250, if they could just find the space and money to build it.

3 Alexander W. Astin, *What Matters in College? Four Critical Years Revisited*. (Jossey-Bass, Inc., San Francisco, CA, 1993), 1st ed.

4 David Johnson, Geoffrey Maruyama, Roger Johnson, and Deborah Nelson, "Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis," Psychological Bulletin **89**, 47-62 (1981); Michael Prince, "Does active learning work? A review of the research," Journal of Engineering Education **93** (3), 223-231 (2004); Lev S. Vygotsky, *Mind in society: The development of higher psychological processes*. (Harvard University Press, Cambridge, MA, 1978).

5 Jack Wilson coined the term "studio physics" for his classroom at Rennsaeler Polytechnic Institute. See J. Wilson, "The CUPLE physics studio," Phys. Teach, 32, 518 (1994). Also see P. Laws, Workshop Physics Activity Guide (John Wiley & Sons, New York, 1997) and "A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms," C. Hoellwarth, M. Moelter, and R. Knight, Am. J. Phys. 74, 374 (2005).

6 Jon Gaffney, Evan Richards, Mary Bridget Kustusch, Lin Ding, and Robert Beichner, "Scaling up education reform," Journal of College Science Teaching **37** (5), 48-53 (2008).

7 Robert Beichner, Jeff Saul, David Abbott, Jeanne Morse, Duane Deardorff, Rhett Allain, Scott Bonham, Melissa Dancy, and John Risley, "Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) project", in *PER-Based Reform in University Physics*, edited by E. F. Redish and P. J. Cooney (American Association of Physics Teachers, College Park, MD, 2007), Vol. 1.

8 Reported at http://scaleup.ncsu.edu/groups/adopters/wiki/48252/Florida_State.html. Accessed 9/9/2008.

9 Richard Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," American Journal of Physics **66** (1), 64-74 (1998).

10 Laird Kramer, Eric Brewe, and George O'Brien, "Improving Physics Education through a Diverse Research and Learning Community at Florida International University", in *APS Forum on Education Newsletter* (American Physical Society, College Park, MD, 2008), pp. 29-32.

11 Danielle Goodwin, internal report found at http://scaleup.ncsu.edu/groups/adopters/wiki/03ed0/Penn_State_Behrend.html. Accessed September 9, 2008.

12 Reported at http://scaleup.ncsu.edu/groups/adopters/wiki/15c1e/Pittsburgh.html. Accessed 9/9/2008.

13 Lin Ding, Ruth Chabay, Bruce Sherwood, and Robert Beichner, "Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment," Physical Review Special Topics - Physics Education Research **2** (010105), 7 (2006).

14 Maria Oliver-Hoyo and DeeDee Allen, "Effects of an Active Environment: Teaching Innovations at a Research I Institution," Journal of Chemical Education **81** (3), 441-448 (2004).
15 Scott Schiff, internal report found at http://www.clemson.edu/ese/research/SCALEUP/scaleCiE.pdf. Accessed Septermber 25, 2008.
16 Personal communication, September 18. 2008.
17 Elizabeth J. Wolfe, "Evaluating the use of the SCALE-UP teaching methodology for an undergraduate database systems course," University of Victoria, 2008.
18 Yehudit Judy Dori and John Belcher, "How Does Technology-Enabled Active Learning Affect Undergraduate Students' Understanding of Electromagnetism Concepts?," Journal of the Learning Sciences **14** (2), 243-279 (2005).
19 Yehudit Judy Dori, Erin Hult, Lori Breslow, and John W. Belcher, "How much have they retained? Making unseen concepts seen in a freshman electromagnetism course at MIT " Journal of Science Education and Technology, **16** (4), 299-323 (2007).
20 David Hestenes, Malcolm Wells, and Gregg Swackhamer, "Force concept inventory," Physics Teacher **30** (3), 141-158 (1992).
21 Ronald K. Thornton and David 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 **66** (4), 338-352 (1998).
22 David P. Maloney, Thomas L. O'Kuma, Curtis J. Hieggelke, and Alan Van Heuvelen, "Surveying students' conceptual knowledge of electricity and magnetism," American Journal of Physics **69** (S1), S12-23 (2001).
23 The ECCE and other assessment instruments are available at the Workshop Physics website, http://physics.dickinson.edu/%7Ewp_web/wp_resources/wp_assessment.html.
24 David Scott Abbott, "Assessing student understanding of measurement and uncertainty," PhD, North Carolina State University, 2003; Duane Deardorff, "Introductory Physics Students' Treatment of Measurement Uncertainty," PhD, North Carolina State University, 2001.
25 Janice Streitmatter, *Toward Gender Equity in the Classroom: Everyday Teachers' Beliefs and Practices*. (State University of New York Press, Albany, NY, 1994).
26 Danielle Goodwin, internal report found at http://scaleup.ncsu.edu/groups/adopters/wiki/03ed0/Penn_State_Behrend.html. Accessed September 9, 2008.
27 Deb Alexander, Bradley A. Cohen, Steve Fitzgerald, Paul Honsey, Linda Jorn, John Knowles, Peter Oberg, Jeremy Todd, J. D. Walker, and Aimee Whiteside, "Active Learning Classrooms Pilot Evaluation", (University of Minnesota, Minneapolis, MN, 2008), pp. 6.
28 Carol L. Colbeck, Alberto F. Cabrera, and Patrick T. Terenzini, "Learning professional confidence: Linking teaching practices, students' self-perceptions, and gender," Review of Higher Education **2**, 173-191 (2001).
29 Arthur Bandura, *Social foundations of thought and action: A social cognitive theory*. (Prentice Hall, Engelwood Cliffs, NJ, 1986).
30 S. Biggers L. Benson, W. Moss, M. Ohland, M. Orr, S. Schiff, "Adapting and Implementing the SCALE-UP Approach in Statics, Dynamics, and Multivariable Calculus", in *ASEE Annual Conference and Exposition* (Honolulu, HI, 2007).
31 Danielle Goodwin, internal report found at http://scaleup.ncsu.edu/groups/adopters/wiki/03ed0/Penn_State_Behrend.html. Accessed September 9, 2008.
32 Standards for several different types of programs are available at http://www.abet.org. Accessed Sept. 10, 2008.
33 R. J. Beichner, "Testing student interpretation of kinematics graphs," American Journal of Physics **62** (8), 750-762 (1994).

34 Paula Vetter Engelhardt and Robert J. Beichner, "Students' understanding of direct current resistive electrical circuits," American Journal of Physics **72** (1), 98-115 (2004).
35 Edward F. Redish, Jeffrey M. Saul, and Richard N. Steinberg, "Student expectations in introductory physics," American Journal of Physics **66** (3), 212-224 (1998).
36 Daniel MacIssac and Kathleen Falconer, "Reforming physics teaching via RTOP," Physics Teacher **40**, 479-485 (2002).
37 Eric Mazur, *Peer Instruction: A User's Manual*. (Prentice-Hall, Upper Saddle River, NJ, 1997).
38 Private communication
39 Linda L. Carli, "Gender and social influence," Journal of Social Issues **57** (4), 725-741 (2001).
40 Charles Henderson and Melissa H. Dancy, "Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics," Physical Review Special Topics - Physics Education Research **3** (2), 020102 (2007).
41 David Pundak and Shmaryahu Rozner, "Empowering engineering college staff to adopt active learning methods," Journal of Science Education and Technology **17** (2), 152-163 (2007).