Title
Teaching and physics education research: bridging the gap

Journal
Reports on Progress in Physics, 77(3)

ISSN
0034-4885

Authors
Fraser, James M
Timan, Anneke L
Miller, Kelly
et al.

Publication Date
2014-03-01

DOI
10.1088/0034-4885/77/3/032401

Key Issues Review

Teaching and physics education research: bridging the gap

James M Fraser[1,2], Anneke L Timan[2], Kelly Miller[1,3], Jason E Dowd[1,3], Laura Tucker[1,3] and Eric Mazur[1]

[1] Department of Physics and School of Engineering and Applied Sciences, Harvard University, 9 Oxford Street, Cambridge, MA 02138, USA
[2] Department of Physics, Engineering Physics and Astronomy, Queen's University, Kingston, Ontario, K7L 3N6, Canada

E-mail: mazur@seas.harvard.edu

Received 21 July 2013, revised 20 December 2013
Accepted for publication 21 January 2014
Published 4 March 2014

Invited by Steve Ritz

Abstract
Physics faculty, experts in evidence-based research, often rely on anecdotal experience to guide their teaching practices. Adoption of research-based instructional strategies is surprisingly low, despite the large body of physics education research (PER) and strong dissemination effort of PER researchers and innovators. Evidence-based PER has validated specific non-traditional teaching practices, but many faculty raise valuable concerns toward their applicability. We address these concerns and identify future studies required to overcome the gap between research and practice.

Keywords: physics education, active learning, problem solving, student resistance, faculty time, research-based instruction, student learning

1. Introduction

Physics faculty—experts in evidence-based research—often rely on anecdotal experience to guide their teaching practices. Adoption of research-based instructional strategies[4] remains surprisingly low, despite a large body of physics education research (PER) spanning over two decades, and extensive dissemination efforts by physics education researchers and innovators (Turpen and Finkelstein 2009, Dancy and Henderson 2010, Turpen and Finkelstein 2010). Traditional lectures (teacher-focused, with passive students) still reign as the instructional strategy of choice in many science courses (Hertado *et al* 2012). The continued prevalence of this approach is surprising given the dramatic gains achieved by research-based instructional strategies: improved conceptual understanding, increased retention in enrollment, and reduction of well-documented minority gaps (e.g., gender, ethnic). These improvements are becoming increasingly important as we face a shortfall of skilled workers in science-related disciplines. The President's Council of Advisors on Science and Technology reported in February 2012 that 'economic projections point to a need for approximately 1 million more (science, technology, engineering, and mathematics) professionals than the US will produce at the current rate over the next decade.' A key problem identified by the report is that 60% of students who enter college intending to major in science-related field do not graduate with a science degree.

Furthermore, dramatic advances in information technology require a re-examination of the basic role of the physics instructor. When huge volumes of information and sophisticated search algorithms are as close as the student's smartphone, the role of instructor as content deliverer seems

[3] These authors contributed equally to this paper.
[4] A research-based instructional strategy is a specific pedagogical approach that has shown effectiveness through empirical measurement (quantitative and/or qualitative). Evaluation often takes the form of pre- and post-testing of student comprehension and skills to evaluate learning, and comparison to a control group that was taught using an alternate technique. The evaluation of research-based instructional strategies in a 'real' setting (as part of a standard course) is particularly valued even though it often does not allow the study design to reach the same level of rigor possible in other fields (e.g., medical trials).

destined for obsolescence. Even the perspective that the in-class instructor will always be needed to organize and explain content is challenged by the recent rise of massively online open courses. These free offerings deliver physics and other subject material in a slick package narrated by an engaging instructor. Large student numbers and flexible administrative structures allow administrators of these courses to redesign them and quantitatively assess changes much more rapidly than they would through a traditional university's curriculum development. Even if the reader is completely confident that massive online open courses can never replace the learning environment made possible through direct human interaction, education funders (upper administration and government officials) will be strongly attracted to the apparently huge 'productivity gains' (number of students 'output' per dollar invested) offered by such an approach. A pessimistic viewpoint would be that with increasing pressure to tighten operating budgets and/or increase enrollments, instructors will be reduced to the role of evaluator, akin to a 'driving test examiner' for university degrees.

As instructors, it is in our best interest to focus on teaching approaches that yield high rates of student success and exploit learning technology to enable these strategies. By taking a scientific approach, PER seeks to identify which instructional techniques work and which do not. The fact that instructors have not embraced research-based instructional strategies as readily as they regularly assimilate novel theoretical or experimental approaches in their technical research field is troubling. We see the problem arising from three sources.

(1) The applied nature of PER, where researchers have incomplete control of all the input experimental variables, is unfamiliar to physics researchers who are used to carefully controlled studies. Results are easily dismissed as being relevant to the specific student population and particular instructor in the study. PER requires validation with many instructors at various institutions before it can provide a clear picture of the general applicability of an instructional strategy. As Prince states, 'The more extensive the data supporting an intervention, the more a teacher's students resemble the test population and the bigger the reported gains, the better the odds are that the method will work for a given instructor' (Prince 2004). By synthesizing the results from many studies, this review article attempts to demonstrate to the non-expert that research-based pedagogy and in particular, interactive engagement, has been strongly validated and gives excellent odds of student success.
(2) PER sometimes overlooks issues essential to an instructor's ability to successfully implement a teaching strategy. It is not surprising that practitioners have limitations that are different from those of physics education researchers (such as time to devote to learning about teaching). We identify specific challenges that require further research so that education researchers can better bridge the gap between research and practice.
(3) Instructors suffer from insufficient feedback to help them improve student learning. While physics researchers benefit from the culture of scholarship in their technical fields (and the external peer review that it entails), such feedback is often lacking in their classrooms. Specific and timely feedback, collected from a sufficient sample of the instructor's students would greatly improve the instructor's motivation and ability to adopt novel instructional strategies.

The target audience for this review paper is physics researchers who want to maximize student learning in their lecture rooms, but may be skeptical about the relevance of PER for their classroom practices. A second target audience is physics education researchers who want to better understand the concerns of front-line instructors to improve the alignment of their research with user needs.

Many education innovators have developed and tested instructional strategies, which range from implementations easily introduced in the large lecture hall to modifications that require extensive redesign of university infrastructure and administrative structure. At the core of most successful implementations are interactive engagement techniques that enable active learning with many opportunities for formative feedback to the student. In his landmark paper, Hake defines interactive engagement methods as 'those designed at least in part to promote conceptual understanding through interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback through discussion with peers and/or instructors' (Hake 1998). Henderson and Dancy provide a detailed list of research-based instructional strategies that have enjoyed considerable success in physics education, such as Peer Instruction, Physlets, Cooperative Group Problem Solving, Workshop Physics, Just in Time Teaching, Tutorials in Introductory Physics, Interactive Lecture Demonstrations, Activity-based Problem Tutorials, Ranking Tasks, SCALE-UP, Active Learning Problem Sheets, RealTime Physics Laboratories, and Context Rich Problems (Henderson and Dancy 2009). For a curated list and related links the reader is referred to the Physics Education Research User's Guide, maintained by the American Association of Physics Teachers (www.compadre.org/perug/). Due to the considerable literature on implementation and verification of these techniques, this review does not attempt to evaluate or identify which approach is best for instructors' particular needs. We see this as a question of less consequence compared to the more important issue: what are the main obstacles to the adoption of research-based instructional strategies? As summarized by the President's Council of Advisors on Science and Technology, 'Based on the weight and variety of the research evidence, it is reasonable to say that if one active-learning event in which students engaged and received feedback were incorporated into each classroom session of every introductory college class in the United States, science education would likely be transformed.'

An emerging consensus in the literature focuses the definition of active learning on engaging every student. We modify Hake's definition to add this important emphasis: interactive engagement methods promote conceptual understanding through interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate individual feedback to all students through discussion with peers and/or instructors. This definition shows that many techniques that appear to increase engagement in a subset of students are not sufficient to be considered truly active learning. Examples of passive teaching techniques masquerading as active learning include: entertaining lectures, demonstrations, multimedia presentations, recipe labs (hands-on, yes, but perhaps not heads-on), class-wide discussion involving a minority of students, and classroom response systems used primarily for attendance taking or testing memory recall. These approaches can be a useful component to an active learning approach but are not sufficient in themselves. A possible litmus test to determine if an instructor is creating an active learning environment is provided by the following question: in every lecture, does every student present an idea, justify it, and evaluate it through critical feedback from colleagues or an instructor?

We have compiled faculty responses to research-based instructional strategies through surveys reported in the literature and feedback from participants who attended one of many colloquia or workshops presented by our group. Common concerns are grouped according to major themes. Each theme starts with typical (paraphrased) concerns expressed by instructors. The themes we address are the following.

(I) Can we approach our teaching with scientific rigor?
(II) What's wrong with the traditional lecture?
(III) Does a focus on conceptual understanding short-change problem-solving skills?
(IV) Do students have time and inclination for research-based instruction?
(V) How will faculty find the time to adopt and maintain research-based pedagogy?

We highlight specific studies and meta-analyses that respond to each theme, and propose new research directions that are required to address unresolved issues. Some studies are relevant to multiple themes and are addressed more than once; each section is intended to be self-contained to allow the reader to consider the themes that resonate most.

Theme I. Can we approach our teaching with scientific rigor?

> '*Teaching is really an art, not a science. Some instructors have the "magic", and some do not.*'

> '*A strategy that works well for one instructor can easily fail with another instructor. There are simply too many uncontrolled parameters to generalize the results of physics education research.*'

Advances in physics occur because physicists rigorously apply the scientific method. It is logical that physics researchers should approach their teaching with similar scientific rigor. However, anecdotal evidence, biased samples, and poor feedback mechanisms often prevail over scientific approaches for evaluating student learning. For example, after providing an elegant derivation or explanation on the blackboard, a professor might ask his students if they understand what he has taught. Observing nods in the front row, the professor might conclude that he has taught the material well. Another professor, known for her charismatic lectures and entertaining demonstrations, might assume that her teaching was successful when she receives glowing student evaluations at the end of the semester. Yet another professor might rely on his own experiences as a student, assuming that his students will learn well from traditional lectures because he learned this way (see theme II for a detailed discussion). In the absence of a scientific approach for teaching evaluation, these biased feedback mechanisms perpetuate the use of less-than-optimal teaching strategies.

PER seeks to apply the scientific method to improve teaching; however, sometimes unfamiliarity with education research methods leads physicists to distrust this type of research. Unlike research conducted within a physics laboratory, where parameters are carefully controlled, PER relies on data from diverse classrooms, yielding variable results. These concerns are important to address. If PER is indeed an effective scientific pursuit, we require a critique of the scientific method used to improve teaching.

A recent study by Deslauriers *et al* (discussed in more detail in theme II) found impressive results of a three hour intervention with over 500 students (Deslauriers *et al* 2011). A control group was taught using a traditional lecture format by an award-winning faculty member, whereas an experimental group was taught using interactive research-based instructional strategies with a relatively novice teaching fellow (and graduate student assistant). Following this brief intervention, the experimental group scored significantly higher on a test of physics conceptual understanding and on other metrics, such as student engagement and attitudes. From these data, the report suggests generalizability to a 'variety of post-secondary courses'.

In the popular press and on website blogs, the study was criticized for the scientific soundness of the researchers' methods. This criticism raised a number of common questions. First, were the instruments for measuring learning valid? In the Deslauriers study, the instructor of the experimental group prepared a multiple-choice test to measure the students' conceptual understanding, instead of using a validated standardized conceptual test. Second, how sure can we be that a study is replicable and generalizable? Deslauriers *et al* did not have a true randomized trial (Derting *et al* 2011), and with only two instructors, the researchers could not control for alternative explanations of their results, such as individual instructor effects (Torgerson 2011). The reactions to the Deslauriers study do not discount its notable results, the important implications, and the widespread impact the study has had across scientific disciplines. Rather, the critiques are a reminder that PER is a science, and should be assessed with scientific rigor. Under themes III and IV, we discuss the importance of validated measuring instruments, generalizability, and replication for PER to be regarded as a useful scientific pursuit.

In our opening illustration, we described three weak feedback mechanisms that are commonly used to evaluate teaching. Nods from a biased selection of students in the front row are a poor indicator of whether the whole class (or even the front row) genuinely understands the material. Similarly, end-of-semester student evaluations are often out-of-context and untimely sources of feedback as described in theme V. Weak measurements are not restricted to these qualitative sources, however. Academic grades and unstandardized exams also provide limited and sometimes misleading information about student learning. High scores on these assessments might indicate a high degree of student learning, but they could also indicate rote pattern recognition—when students have practiced solving specific problems rather than learning to be problem solvers. The challenges of accurately measuring problem-solving ability will be further addressed under theme III.

Though no single test can measure all facets of learning, physics education researchers have developed validated methods for measuring specific aspects of physics conceptual understanding, problem-solving abilities, and student attitudes and approaches to learning physics. The Force Concept Inventory (Hestenes *et al* 1992), Force and Motion Concept Evaluation (Thornton and Sokoloff 1998), Conceptual Survey of Electricity and Magnetism (Maloney *et al* 2001), and Brief Electricity and Magnetism Assessment (Ding *et al* 2006) are examples of standardized validated measures of conceptual understanding of first-year mechanics and electromagnetism[5]. These tests allow physics education researchers to compare interventions across different populations in order to identify trends for successful learning gains. Conceptual understanding tests are also useful to physics instructors who want to know how their students are thinking about the concepts they teach. Results from these validated instruments can stimulate change; if instructors quantitatively observe that their own students are only learning a small fraction of what they thought they were teaching, these instructors may be more motivated to experiment with different teaching practices.

Applying a postmodern framework, in which quantitative descriptors are simply one way of understanding how people learn, qualitative research can also provide the rich descriptions sometimes lacking in quantitative research (Denzin and Lincoln 2005). It is important to note that ‘qualitative’ should not be confused with ‘subjective’. Examples of qualitative measures include coding student descriptions of their problem-solving strategies and analyzing student interview responses (Hegde and Meera 2012). Qualitative analyses of videotaped problem-solving sessions (Bing and Redish 2009) can be used in conjunction with quantitative validated tests of problem-solving ability (Marx and Cummings 2010) to better understand how students approach physics problems. Students’ attitudes and approach to learning physics can be measured qualitatively or quantitatively through validated surveys such as the Colorado Learning Attitudes about Science Survey (Adams *et al* 2006); this validated survey was successfully used in the Deslauriers study. PER, like all scientific fields, requires agreement across a variety of analysis techniques in order to draw meaningful conclusions.

While a single physics education study provides clues regarding the effectiveness of a teaching strategy, we require replication and meta-analysis before the pedagogy can be generalized to a variety of classrooms. If performing an education study is like forming a steel beam, then meta-analysis is the construction of a bridge out of those beams—identifying the pedagogies most likely to yield success across varied classroom environments (Cooper *et al* 2009). Physics instructors can be increasingly confident in applying a teaching strategy as experiments are successfully replicated and analyzed in context with related PER. This meta-analysis requires strong citations; when current research builds on past discoveries, consistently successful teaching trends can be identified.

Effective meta-analysis also requires researchers to report ‘failed’ experiments in which the novel teaching method does not yield the hypothesized learning gains. For example, one physics education study compared a new Matter and Interactions experimental course to a traditionally taught course. These researchers found gains[6] of 0.21 in the experimental course versus 0.43 in the traditional course: the intervention did not improve students’ conceptual understanding (Caballero 2011). Physics education researchers need to report both positive and negative findings for effective meta-analysis. Just as the first professor in our opening illustration should not judge student understanding on the biased sample of positive nods from front-row students, physics education researchers should not base their claims on positive results from a few ‘front row’ studies. Publishing unsuccessful experiments along with the successful builds a broader understanding of the trends in instructional strategies that tend to accomplish certain learning goals, and those that tend not to.

The strength of PER lies not in one-off studies, but in the assembly and analysis of multiple studies. Many meta-analysis studies have coalesced research on solutions to specific challenges in physics education. The aforementioned meta-analysis by Hake demonstrated a consistent trend: students involved in active learning outperform those who learn passively in a traditional lecture setting (Hake 1998). Another meta-analysis of positive student–teacher relationships—where students are trusted with responsibilities, shown empathy, and spoken to honestly—found correlations between these learner-centered positive relationships and successful cognitive outcomes for students (Cornelius-White 2007). In another example, student collaboration in small learning groups was found to promote academic achievement and positive attitudes toward learning science (Springer *et al* 1999). Researchers continue to combine education studies to address a wide variety of issues, from the effectiveness of distance education (Bernard *et al* 2004) to the gender gap in science (Weinburgh 1995). Through meta-analyses such as these, we can determine the extent to which a positive or negative result can be attributed to the tested instructional strategy or other variables, such as class demographics or instructor effects. These processes allow researchers to identify trends and reduce the impact of the uncontrolled variables that arise in PER. It is in the best interest of physics instructors to make use of these thorough analyses to select the teaching methods most likely to yield the highest learning gains in their classrooms.

[5] For a large listing of inventories, see www.ncsu.edu/PER/TestInfo.html.

[6] Normalized gain refers to the comparison between an exam given as a pretest and posttest. Normalized gain is the ratio of the gain over the maximum possible gain (Hake 1998).

Research-based instructional strategies do not guarantee success in every classroom. However, PER does use an effective scientific approach to discover which teaching strategies provide the best odds for maximum student learning. Just as a sports fan would bet his/her money on the team with the best odds of winning, PER provides instructors with the information they need to choose teaching strategies that are most likely to promote student success. Physics education researchers need to replicate prior research using validated learning measurement instruments, reporting both successful and unsuccessful interventions, so that meta-analysis can reduce the influence of uncontrolled variables and increase the generalizability of instructional strategies. Since PER is a scientific pursuit, both producers and consumers of research-based pedagogy should approach their teaching and teaching research with scientific rigor.

Theme II. What's wrong with the traditional lecture?

> '*I learned well from a clear, logical, and well-explained lecture, so this is the experience that I strive to provide my students.*'

> '*I need to cover a lot of content in the syllabus, and traditional lecture is the most efficient way to do that. Students need me to explain the material so that they can learn it.*'

The majority of instructors believe that the best method for training scientists is the traditional lecture. In 2011, the Mathematical Association of America surveyed over 700 calculus instructors on what they believed was the most effective way to teach. Two-thirds of those surveyed agree with the statement: 'Calculus students learn best from lectures, provided they are clear and well-prepared' (Bressoud 2011). There is a common belief in teaching that, because the instructor is the expert in the room, it his/her job to take the lead role in guiding students through the material. Instructors with this traditional view would argue that because students are novices, active engagement techniques where students are 'teaching' each other could only lead to an increase in confusion about the concepts.

In addition to serving as the expert, most instructors also believe that their role is to meet their students' expectations in the classroom. Students can be resistant, at least at first, to new instructional techniques. If students want informative and well-structured lectures, then the traditional lecture meets students' expectations by allowing them to passively sit, listen, and take notes. To summarize, even though both students and instructors have considerable experience in teaching and learning, this does not directly translate to them being experts in the assessment of various teaching practices.

Science education literature has dozens of studies that quantitatively compare student learning in traditional classrooms to learning in classrooms that employ interactive teaching methods. The bulk of this research indicates that students taught interactively have considerably better conceptual understanding and problem-solving skills than students taught with lecture (Thacker *et al* 1994, Hake 1998, Crouch and Mazur 2001, Meltzer and Manivannan 2002). Furthermore, active learning correlates with better attendance, improved student 'affect'[7] and higher retention in the sciences Adams *et al* 2006, Brewe *et al* 2009, Watkins and Mazur 2013, Deslauriers *et al* 2011).

One of the most highly cited studies to compare student conceptual learning in traditionally taught environments to interactive classrooms was a meta-analysis conducted by Hake (1998), as mentioned in theme I. Hake used six-thousand students' normalized gain scores[8] on the Force Concept Inventory to compare the effectiveness of traditional instruction and of interactive engagement techniques. He found that the interactive courses are, on average, more than twice as effective in producing basic conceptual understanding as traditional courses. The fourteen traditional courses in Hake's study, containing a total of 2084 students, achieve a normalized gain of 0.23 ± 0.04 compared to the 48 interactive engagement classes, containing a total of 4458 students, which achieve an average gain of 0.48 ± 0.14 (Hake 1998). Crouch and Mazur found very similar differences in conceptual learning between students taught traditionally and those taught using Peer Instruction (Crouch and Mazur 2001). In another study, Meltzer investigated the effectiveness of an active learning strategy over a seven-year period at four large universities. Meltzer found that student conceptual gains, measured with the Conceptual Survey in Electricity and Magnetism, are three times greater for students in an interactive lecture than the average gain from a national sample of classes that used a traditional lecture strategy (Meltzer and Manivannan 2002). Alongside such conceptual gains, the literature also indicates that, in spite of a shift away from explicit problem solving in research-based instructional strategies, improvements in students' ability to solve problems are at least comparable to those achieved through traditional lecture. Theme III discusses this point in detail.

Physics education literature also reveals that students taught interactively have a higher level of engagement and are more likely to persist in science-related fields than those taught traditionally. Recently, the Carl Wieman Science Education Initiative conducted a controlled experiment comparing two sections of a class taught using different strategies (Deslauriers *et al* 2011). During the first half of the semester, both sections were taught using traditional lectures. Halfway through the semester, the teaching strategy for one of the sections was switched to a design informed by 'deliberate practice' (Anders Ericsson *et al* 1993) involving many aspects of interactive engagement for a period of one week. Students' attitudes toward science and level of engagement were measured using class attendance rates and the Colorado Learning Attitudes about Science Survey (CLASS). Trained observers also watched both sections to measure student engagement. Before the experiment (halfway through the semester), both sections were statistically equivalent with

[7] Affect is an encompassing term, used to describe the topics of emotion, feelings, and moods together; it is commonly used interchangeably with emotion (Finkelstein 2013).

[8] Gain is the difference between students' before- and after-instruction scores on an exam. Normalized gain is the ratio of the actual gain over the maximum possible gain (Hake 1998).

respect to knowledge, attendance, and level of engagement. At the end of the one week test period, results showed that while engagement and attendance for the control section remained unchanged, in the interactive section, student engagement doubled and attendance increased by 20%.

Other studies (Adams *et al* 2006, Brewe *et al* 2009) have demonstrated that students' attitudes about physics and physics learning improve in introductory courses using Modeling (Wells *et al* 1995). The Modeling Method addresses many of the drawbacks of the traditional lecture by engaging students in activities that help them organize concepts into viable scientific models. Most importantly, this improvement in student attitude has been found to influence career choices in science-related fields. Watkins found a positive relationship between interactive teaching and persistence in the sciences (Watkins and Mazur 2013). Regardless of the race or gender of the student, Watkins found that those enrolled in just one interactively taught introductory physics class are twice as likely to pursue science as their major compared to those enrolled in only traditionally taught classes. Given the current and future shortage of science professionals discussed in the introduction of this paper, the positive correlation between interactive teaching and student persistence could be considered even more important than conceptual understanding gains.

Despite many instructors' belief that students learn best by listening to expert lectures, the literature demonstrates that this is not the case. The evidence for significantly higher conceptual understanding gains in active learning environments over passive ones is hard to overlook. Anecdotally, one can imagine why Peer Instruction and other interactive engagement strategies are so effective. As novices, students speak the same language, while most instructors, as experts, have long since lost the ability to speak like a novice. As a result, fellow students can play an essential teaching role in partnership with instructors. Furthermore, students are more likely to be critical of explanations provided by peers than those provided by faculty who they perceive as experts. This critical listening promotes learning and helps novices transition to more expert ways of thinking. In fact, a large body of data indicates that students taught interactively are more engaged and persistent in the pursuit of further study in science-related fields. These results show that, on average, active learning strategies strongly benefit students in both the short and long term.

Theme III. Does a focus on conceptual understanding short-change problem-solving skills?

> '*Active learning strategies focus too much on basic conceptual understanding. My students need problem-solving skills to become good physicists.*'

> '*If I don't show my students how I solve problems on the blackboard, they won't know how to solve problems on their own.*'

Perhaps it seems strange that in PER, the litmus test for effective instruction often is the demonstration of students' conceptual understanding rather than problem-solving skills. After all, quantitative analysis of the physical world is the foundation of the discipline of physics. Decades of research show that experts and novices solve problems differently (Larkin and Reif 1979, Larkin *et al* 1980, Chi *et al* 1982, Smith 1991, Priest and Lindsay 1992), and that differences in strategy are related to differences in performance. Instructors of intermediate and advanced courses expect students to have developed problem-solving skills at the introductory level. So one might wonder why PER focuses on improved conceptual understanding, rather than on physics problem-solving strategies.

Several studies have, in fact, focused squarely on the development and analysis of classroom interventions to improve students' problem-solving skills. Interventions range from carefully crafted heuristic scaffolding (Heller and Reif 1984, Wright and Williams 1986) to worked examples (Ward and Sweller 1990) and categorization activities (Mestre *et al* 1993) to entirely redesigned curricula (Heuvelen 1991a, 1991b, Heller and Hollabaugh 1992, Heller *et al* 1992, Leonard *et al* 1996). In each case, students who participated in the interventions outperformed students who did not participate in them. A meta-analysis of 22 studies indicates several shared features—heuristic scaffolding, modeling by the instructor, and explicit requirement that students follow the procedure—each contribute to improved problem-solving performance (Taconis *et al* 2001). If improved problem solving is the desired outcome, there is evidence to support particular approaches to instruction.

While these studies were carried out during the middle 1980s and early 1990s, it was generally believed that strong problem-solving skills correspond to strong understanding. During this same period of time, however, research into students' conceptual understanding was beginning to show that students often do not understand many basic concepts even at the end of instruction and in spite of an ability to solve textbook problems (Halloun and Hestenes 1985, Hestenes *et al* 1992, Crouch and Mazur 2001). The alarming lack of conceptual understanding prompted investigators to shift the basis of comparative analyses of instructional interventions from the latter to the former.

Over subsequent years, the importance of improving stubbornly poor conceptual performance became a primary focal point of PER. As novel research-based instructional strategies embraced the new focus (Hake 1998), explicit reference to their effect on students' problem-solving ability drew less attention. A relatively small fraction of the studies published in the late 1990s and 2000s draw explicit attention to students' problem-solving skills in conceptually focused instructional strategies; in fact, many studies press even further outward with questions of students' beliefs about physics (Adams *et al* 2006, Brewe *et al* 2009), sense of identity (Hazari *et al* 2007, Hazari *et al* 2010) and participation in a physics community (Brewe *et al* 2012).

Ultimately, if the goal is to improve education and the largest deficiencies are observed in conceptual understanding, then improving conceptual understanding is the most natural place to start. A few central concepts, properly understood, provide the same explanatory power as a plethora of equations. However, problem-solving skills must not be neglected; students must be able to reason with graphs and numbers. Just as one should not assume conceptual understanding follows from enhanced problem solving, one should not assume that improved performance on conceptual surveys is linked to an improved ability to solve problems.

It is important to recognize that our understanding of the nature of problem solving is changing even as investigators attempt to assess students' problem-solving ability. Many studies, including some described here, tend to treat problem-solving ability as an unambiguous construct. However, the decades of research in problem solving indicate that it is not so simple.

To some extent, conventional quantitative problems in physics can operate like puzzles (Chi *et al* 1982). Solvers use rules to turn an initial state into a final state—perhaps similar to navigating a maze. Novices tend to rely heavily on a 'means-end' approach, in which one tries to connect the final state to things that are known about the initial state. In contrast, experts tend to work forward and contextualize the situation according to physical models. Solutions to traditional homework, textbook, and exam questions tend to fit into these paradigms.

However, questions—in which the procedure is more-or-less known and only the answer is missing—differ from actual problems—in which the answer is often known, but the procedure is missing. Some of the early investigators of students' quantitative reasoning found that students attempt to subvert the designed heuristics or fail to make the desired connections among concepts (Heller and Hollabaugh 1992, Heller *et al* 1992, Leonard *et al* 1996). In response, these researchers redesigned the problem-solving activities to incorporate context-rich and real-world qualities, resulting in more authentic measures of students' quantitative problem-solving abilities.

Despite the limited scope and educational value of conventional physics problems, students' ability to solve these types of problems is often valued by instructors. Future studies certainly must incorporate more broad and varied notions of problem solving, but there remains great value in discussing the work that has been done in evaluating the effectiveness of various teaching techniques on conventional problem-solving skills.

In 1994, Thacker and colleagues compared students' problem-solving performance in concept-focused and traditional introductory physics courses (Thacker *et al* 1994). Students enrolled in the concept-focused class were elementary education majors, while students enrolled in the three traditional classes were non-scientists, engineering students, and physics majors, respectively. In this study, two problems—one conceptual and one quantitative—were assessed on midterm or final exams. Students in the concept-focused class performed best of all classes on the conceptual problem and second only to students in the honors class on the quantitative problem. The authors argue that concept-focused students perform as well as honors students by treating the sum of the two questions as performance, though they could also have emphasized the fact that conceptually focused students are at least as strong at quantitative problems as their counterparts in traditional courses. Differences in the incorporation of course content and the questions asked in each class certainly played a role in this finding. Nonetheless, the study highlights the fact that elementary-level educators, who are typically non-scientists with a limited background in physics, can perform quite well on quantitative questions after concept-focused instruction.

Subsequent studies involving different research-based instructional strategies support similar conclusions; this replication is crucial for generalizing these results, as described in theme I. In 2000, Jones and colleagues reported that students in a concept-focused interactive class perform better than students in a traditional course on conceptual post-course surveys, three conceptual exam questions, and two out of three quantitative exam questions (Jones *et al* 2000). Crouch and Mazur found that students in Peer Instruction classes exhibit larger gains (see footnote 7) on both the Mechanics Baseline Test, which includes both qualitative and quantitative questions, as well as the subset of quantitative questions from that test (Crouch and Mazur 2001). They also found that students in Peer Instruction performed better in exam situations—one year involving an entire quantitative exam, and another year involving a single shared quantitative question on the final exam—than students in a traditional class. Upon implementing Peer Instruction at a two-year college, Lasry and colleagues reported no significant differences in final exam performance between interactively- and traditionally-instructed students (Lasry *et al* 2008). Meltzer and Manivannan reported that students in their interactive classes outperformed students in traditional, calculus-based classes on shared final exam problems (Meltzer and Manivannan 2002). Somewhat contrary to these findings, Hoellwarth and colleagues also compared shared final exam problems between students in an interactive learning environment and those in a traditional learning environment, and found that the latter group outperformed the former group in one out of three years (Hoellwarth *et al* 2005). The differences between the two groups are not statistically significant during the other two years, and conceptual gains are always stronger among those in the interactive environment.

These results suggest that concept-focused, research-based instructional strategies can improve students' problem-solving skills as effectively as traditional approaches to instruction, even though such strategies often involve much less class time devoted to students observing an expert solve typical problems. Interestingly, the interactive learning environment of Hoellwarth and colleagues, in which students appear to struggle more with problem solving, involves more in-class, expert-led problem solving than the interactive environment of Crouch and Mazur, in which students appear to struggle less with problem solving. Of course, numerous other differences—instructor, institution, student population, additional variations in pedagogy—may be relevant as well. Without more extensive analysis of a greater pool of studies, one can only speculate about variations within and among the studies highlighted here. Nonetheless, it is consistently evident that quantitative problem-solving skills do not suffer nearly as much under concept-focused pedagogy as conceptual understanding suffers under the traditional approach.

Along with differences in structure among these pedagogies, time spent on various tasks is a particularly important aspect of these results. When students are not offered recitation sections, they are dissatisfied with their preparation for quantitative problems encountered on exams (Jones *et al* 2000). Surveys indicate that students in the conceptual class feel they need to do more work outside of class and are frustrated by not knowing what to expect. This finding motivated the instructors to implement weekly 'supplemental instruction' sections. When homework and examinations problems are mostly quantitative while the rest of the course is not, students depend heavily on access to teaching assistants during weekly office hours (Crouch and Mazur 2001). Some argue that the additional recitation sections and time-on-task required in inquiry-based classes should not be 'controlled for' or treated as separate from the pedagogy because the time-on-task is, indeed, part of the pedagogy (Thacker *et al* 1994). While this perspective may be valid, it is of little consolation to instructors who simply do not have the extra time.

When considering time, one must emphasize how that time is being spent. Evidence suggests that coverage of material does not necessarily result in learning, and students ultimately retain only a fraction of the content delivered (Halloun and Hestenes 1985, Heckler and Sayre 2010, Sayre *et al* 2012). On the other hand, a focus on conceptual understanding tends to enhance retention (Heuvelen 1991b, Leonard *et al* 1996). Ultimately, it is a disservice to the students to demand that they cover more material than they are empirically shown to be able to learn. For a more thorough discussion of students' time, please see theme IV.

Students' ability to solve problems is an important dimension of learning physics. However, as we measure the efficacy of physics education, focusing solely on problem-solving ability is insufficient. Not only is the notion of problem-solving ability an evolving construct, but deficiencies in students' conceptual reasoning are even more dramatic and require attention as well. To the extent that students' problem-solving abilities have been investigated in research-based pedagogy, evidence strongly suggests that problem-solving skills do not suffer nearly as much as conceptual understanding suffers under traditional instruction. In many cases, students tend to solve problems just as well in both learning environments. Finally, although constraints to students' time may seem to limit instructors' flexibility in addressing conceptual content and problem solving, coverage does not equal learning. Therefore, there is value in making time for both aspects of learning physics.

We are well positioned to ask where to go from here. As discussed in other themes, Hake's 1998 meta-analysis of six-thousand students' conceptual understanding under traditional and interactive instructional strategies is one of the most often-cited and widely disseminated publications in PER. Many individual data were collected and analyzed together, as described in theme II, allowing for a strong picture of the relationship between pedagogy and conceptual understanding to emerge. At this time, a similar study has yet to be published regarding pedagogy and problem-solving skills. Moreover, as we continue to develop new, categorically distinct kinds of physics problems for students to solve, we must assess both students' ability to solve such problems (which will necessitate qualitative and quantitative analytical techniques) and how such problem-solving abilities relate to other problem-solving abilities. In other words, we need to build a more systematic, multi-faceted understanding of problem solving. Such work will drive research-based pedagogy to respond to the specific needs and constraints of physics educators, thereby providing better resources for effective teaching of problem-solving skills.

Theme IV. Do students have time and inclination for research-based instruction?

> '*A full load of interactive courses requires more time than students have, so not all courses should be taught using research-based instructional strategies.*'

> '*Students expect the lecture, plan their schedules around a lecture, and know how to learn from a lecture. It isn't fair to students to ask them to change their way of learning while they are under pressure to learn difficult material.*'

> '*Students don't like research-based instructional strategies.*'

For college students, as for their professors, multiple activities compete for limited time. Students take full course loads, work part- or full-time, participate in extra-curricular activities, and balance family responsibilities. Given that the majority of college courses use traditional lecturing, students plan their schedules around attending lectures and preparing for traditional assessments. Do research-based instructional strategies increase students' time investment in the course? Do students revolt when research-based instructional strategies are implemented? If active learning does take more student time, is that time the main reason research-based strategies achieve higher gains?

In discussing student time use, we distinguish between time use in class and outside of class. In-class time use includes the number of class meetings covering a particular topic and the length of each class. Outside-of-class time use includes the amount of time spent studying the subject excluding time spent in class. Outside-of-class time comprises, for example, homework, reading, memorization, and self-testing, and the specific approach to completing study activities. We first discuss results regarding in-class time comparisons.

A handful of studies that compare research-based teaching to traditional instruction explicitly report holding class time constant and show greater learning gains with equal class time. For example, Redish *et al* compare five control courses using recitation sections with five experimental courses that replaced recitations with interactive engagement, small-group tutorials (Redish *et al* 1997). The tutorial courses showed double the gains (see footnote 7) on a standardized concept test compared to the control courses, with no change in required class time. Similarly, Deslauriers and Wieman's quantum mechanics curriculum using clicker questions with discussion and small group activities not only had significantly higher scores on an end-of-semester Quantum Mechanics Concept Survey than a traditionally-taught comparison using the same amount of class time, but these gains persisted months after the courses finished (Deslauriers and Wieman 2011). Beyond this, Hake's landmark paper of 6000 students and from 62 courses, discussed in previous themes, compares normalized gain on the Force Concept Inventory in interactively and traditionally taught courses (Hake 1998). Hake asserts, 'the gain difference is not very sensitive to the fraction of the course time spent on mechanics over the range common in introductory courses'. Use of interactive teaching, however, was shown to have a significantly positive effect on gain between the beginning and end of the term.

Even without increasing required class time, students in courses using research-based instructional strategies are likely to increase time-on-task through higher attendance rates. Knight and Wood mention increased class attendance in their reformed upper-division biology class, with attendance ranging between 60% and 70% for traditional instruction but over 90% in the interactive course (Knight and Wood 2005). This effect is also seen by Deslauriers *et al* in an introductory physics course (Deslauriers *et al* 2011). In Knight and Wood's study, points were assigned to participation and correct answers, but no points were assigned to participation in the Deslauriers *et al* study. Both increased motivation and motivation for in-class points may drive higher attendance in courses using research-based instructional strategies. In either case, research-based instructional strategies appear to increase in-class time by motivating students to attend, which we see as a success.

We now move to work on student time spent outside of class. There are two common methods to measure time-on-task: student-kept time diaries and student survey self-reports (Kember *et al* 1995). We will briefly explain each method.

Time diary studies ask students to document their studying on a regular basis, often including the specific study activity (Kember *et al* 1995, Zuriff 2003, Ruiz-Gallardo *et al* 2011, Stewart *et al* 2012). Beyond better estimates of student study time, time diaries can give insight into what students focus on during their study time, students' study approach in research-based and traditionally taught courses, and whether students are multi-tasking while trying to study. For example, recent work has shown that it is not uncommon for students to multi-task while doing academic work (Mokhtari *et al* 2009), and this is an important aspect of study time that is not uncovered by simple time measurements. Also, the same study reports students spending nearly 2.5 h daily on the internet, which is more than the 2.2 h spent daily on academic reading. Time diaries would be useful to tease out whether this online time reflects studying, leisure, or multi-tasking. More studies including time diaries would benefit the physics education community, specifically studies that could bridge multiple instructors and teaching strategies.

In student survey self-reports, students are asked to estimate the amount of time spent on studying. These reports are usually collected once for the term, at the end of the term. Some self-reports also ask students questions about time use (Bol *et al* 1999); one such survey has been designed for undergraduate physics courses (Stewart *et al* 2012). Student survey self-reports of time use have the advantage of relative ease to collect. End-of-semester self-reports are often collected automatically by university in student course assessments and could be used to average time commitment for different teaching strategies over instructors, courses, and institutions. It would be useful to know if specific teaching strategies correspond to the study times that instructors target for their students, or whether some teaching strategies correspond to significantly more or less time than instructors expect. However, a caution remains that retrospective distortion may cause these self-reported time use data to be inaccurate. End-of-semester course evaluations of student self-reporting of study time may need to be viewed as an upper-limit in light of Zuriff's evidence that these estimates may be inflated up to 40%, possibly because students are filling out these surveys near the end of the term (Zuriff 2003). Near final exams, when study time may increase or high stress may color perception, students may inaccurately report their semester-long study time based on this short-term memory (Bolger *et al* 2003).

The physics education field does not yet have a large-scale comparison across multiple instructors and institutions of the time spent by students in traditional courses versus courses using research-based instructional strategies. These data are important because students' study time likely depends both on the teaching strategy itself and also the implementation. Turpen and Finkelstein report that implementation of research-based instructional strategies varies significantly between instructors (Turpen and Finkelstein 2009), and it is possible that the same teaching strategy implemented differently would result in different student study time. Beyond variation between instructors, there is evidence that student time requirements may decrease once a reformed course becomes established. This is because as a course develops, instructors revise assignments based on observations and student feedback in previous implementations. This process occurred in a documented, new implementation of problem-based learning in biology (Ruiz-Gallardo *et al* 2011). In its first implementation, students spent 266% more time than instructors had allotted for the subject, but subsequent efforts successfully reduced the student assignment load to expected levels. Given that one-third of physics instructors who adopt research-based instructional strategies discontinue their use, many citing student complaints (Henderson *et al* 2012), this result about student time investment may encourage some faculty to continue their research-based teaching reforms.

In the absence of larger studies comparing student use of study time in science courses, we first discuss smaller studies that include student self-reports of time use. Then we highlight information presented in the broader literature on the impact of research-based instructional strategies on student motivation and how students allocate their study time.

Some studies of courses using research-based instruction mention students' self-reported time-on-task, and the results show that research-based instructional strategies do not necessarily increase students' study time. For example, Boyle and Nicol report that for more than 75% of students in the study, time spent in their course using research-based instructional strategies is the same or less than time reported for traditional courses (Boyle and Nicol 2003). Likewise, when Knight and Wood replaced lectures in an upper-division biology course with an interactive format including cooperative problem solving, group discussion, and frequent in-class formative assessments, they found significantly higher learning gains while student evaluations showed no difference in course workload when varying teaching strategy[9] (Knight and Wood 2005).

Boyle and Nicol's study also shows that some students spend more study time in a course using research-based instructional strategies than a traditional course. This, however, may not be because of increased workload, but rather because students choose to spend more time because they find the activities valuable. For example, in a course using clickers, small-group tutorials, and revised homework, students rate these reformed aspects of the course with the highest utility. Students in these courses report spending seven to nine hours per week on homework, compared to three to four hours per week reported by students in the same course taught traditionally (Chasteen *et al* 2010). The higher study time may not reflect requirements, but choice because students see the assignments as valuable. As with increased class attendance, we see an increase in motivation to spend time studying as a success of research-based instructional strategies. This is particularly relevant in light of Nonis and Hudson's result that mathematics students' time diaries show they have ample time outside of class to study the amount of time expected by instructors, but students choose to spend less time studying than instructors expect (Nonis and Hudson 2010).

While some research-based instructional strategies involve homework not used in traditional instruction, and these new assignments such as tutorial homework (McDermott and Shaffer 2001) or Just in Time Teaching reading questions (Novak *et al* 1999) could increase student workload, the addition of tasks alone is not sufficient to conclude that these teaching strategies increase total time-on-task. It is possible that encouraging students to read the material and answer questions before class gives students a conceptual basis to better absorb material during class time and approach problem sets more efficiently. One example of this effect is found for psychology students who engage in an analysis activity before hearing a lecture (Schwartz and Bransford 1998). These students do better than students who summarize text passages before hearing the lecture, though time-on-task is equal for both groups. Lastly, possibly in agreement, Stewart *et al* found that students' additional time input for new assignments does not scale: when the length of homework and reading assignments increased in their physics course, students' self-reported study time did increase, but the increase was slower than a linear extrapolation based on the number of characters in the reading assignment or number of steps required to solve the homework problems (Stewart *et al* 2012). This could be because students choose to spend a fixed amount of time on their studies and, when assignments increase, students respond by shortening the time spent on any single course aspect. On the other hand, this nonlinear time increase could be because the additional work scaffolds students' learning in a way that helps them to complete additional problems in the assignment more quickly.

Beyond Shwartz and Bransford's work with psychology students showing that well-designed assignments can have learning gains that do not result from merely increasing time on task, we cite a study in introductory physics comparing students' performance on Peer Instruction questions. To compare equivalent time-on-task and the effects on one research-based instructional strategy, Lasry *et al* replaced the peer discussion portion of the standard Peer Instruction protocol with either an individual reflection or a distraction task, all of equal time (Lasry *et al* 2009). They found that the discussion group has far greater learning gains on individual Peer Instruction questions than either the individual reflection or distraction group. These data suggest that the gains in Peer Instruction cannot be attributed to more time-on-task, as greater performance gains are achieved with equal time-on-task in the discussion group compared to the individual reflection group. A second study links learning gains with gains during Peer Instruction. Smith *et al* show in an undergraduate genetics course that gains in correct answer after peer discussions demonstrate learning, as students, after discussion, individually answer a similar question on the same topic with a higher rate of correctness than on their individual answer to the first question (Smith *et al* 2011). When combined, the results of these studies show that even when controlling for time-on-task, interactive activities can still yield higher performance gains.

In addition to task order promoting effective use of student time, frequent formative assessment, an essential component of research-based instructional strategies, may change how student time is spent. One possibility is that frequent assessment helps students who would otherwise do poorly in the course increase their performance. Stewart *et al* report this result, though with summative assessment, for one introductory physics course (Stewart *et al* 2012). In many traditional courses, the first substantive feedback students receive is through exams, which may occur as late as halfway through the term. Earlier assessment allows these students a chance to alter their study approach earlier in the term, perhaps providing another reason for increased retention in some courses using research-based instructional strategies (Lasry *et al* 2008, Brewe *et al* 2010). However, the direct impact of formative assessment on students' time use in physics warrants further study.

Indirectly, research-based instructional strategies may catalyze a change in students' epistemology, thus impacting students' approach to studying. This would be an important result of research-based instructional strategies, and some studies suggest that students in research-based courses do

[9] The authors do mention a possible ceiling effect, as about two-thirds of students rate the workload as too high in both forms of the course. In a ceiling effect, data points cluster at the high end of the measurement scale so that if variations exist between data points at the top of the scale, these variations are not evident on the chosen scale (Hessling *et al* 2004).

approach study time differently. In agreement with this explanation, Stelzer *et al* held required time-on-task constant across control and reformed implementations by replacing some class time with online pre-lectures (Stelzer *et al* 2010). This experiment and its outcome have been reproduced (Sadaghiani 2011). In the experimental sections, students' average difficulty rating of the course decreased even while exam performance increased, suggesting that the students were approaching the material differently in the reformed course. Mauk and Hingley report a quote by one student in an introductory physics class using the *Tutorials in Introductory Physics* that may explain Stelzer's result: 'We now spend the time trying to figure out how the physics works instead of doing calculations' (Mauk and Hingley 2005). Furthermore, students' attitudes toward course content shape their study strategy. Thambiratnam *et al* note that students are more likely to use a surface approach to studying when they see course content as irrelevant (Thambiratnam *et al* 1994). Assessing the value students place on a learning activity may serve as a rough proxy for their studying approach, and also suggests that one mechanism to encourage a deep learning approach is to help students understand the utility of their assignments.

Also, assessment plays an important role in encouraging students to spend their time in a deep learning approach. If assessments center on problems for which memorization of procedures yields success for students, students will have difficulty balancing the goals of high grades and deep learning. Elby's survey of 106 introductory college students in classes using research-based instructional strategies supports this idea. His study finds that while students do not recommend a surface study approach for deep understanding, both high and low-achieving students 'spend more time focusing on formulas and practice problems' than they would recommend for a hypothetical students whose goal is deep learning (Elby 1999). The reason: they perceive that rote (surface) study techniques will earn them good grades. Elby suggests 'less rote-able' homework and exam problems as a potential solution. Further studies would benefit the community to determine rigorously how teaching strategies, including assessments, affect students' studying approach outside of class in both traditional and research-based physics courses. Time diary studies would be particularly useful for this purpose.

We now address student response to research-based instructional strategies. Do students support reformed teaching methods? While most practitioners are familiar with anecdotal evidence of students strongly in favor or against research-based instructional strategies, large-scale studies on the topic show favorable responses by the majority of students. Gasiewski's study of engagement in introductory science and science-related courses surveyed 2873 introductory science students across 15 institutions and 73 courses (Gasiewski *et al* 2011). They found that 'the majority of students preferred [class discussion and group work] and recognized their benefits.' This result strongly supports the idea that most college students appreciate collaboration, a hallmark of most research-based instructional strategies. In agreement of this result, Judson and Sawada's review of electronic response systems over thirty years shows that 'students will favor the use of electronic response systems no matter the nature of the underlying pedagogy'[10] (Judson and Sawada 2002). Learning gains, however, were only associated systematically with the use of these devices when used in conjunction with student discussions.

It is important to note that, while the majority of student responses are positive about the utility of interactive learning, student acceptance of research-based instructional strategies is not uniformly achieved. We encourage instructors who feel student pressure against reformed teaching to gather a representative sample of responses from their students and not be discouraged by a negative, vocal minority. Also, by taking time to help students understand the reasons behind assignments and teaching methods, instructors may gain better buy-in from students. Still, some research-based instructional strategies may be better received than others. Though many studies report high student satisfaction with clickers, one report shows moderate student ratings of utility but consistently low ratings of student enjoyment of University of Washington tutorials (Turpen *et al* 2009). Also, even though support for clickers is largely positive, Caldwell catalogs some recurring student complaints (Caldwell 2007). A large-scale analysis of student response of both utility and enjoyment across multiple instructors and institutions, keeping track of specific teaching strategies used in each classroom, would be an important service to the education research community as well as for physics instructors both to better understand student response and address student complaints.

We find that students' time-on-task can remain the same or even decrease in courses using research-based instructional strategies. Some evidence suggests that students in these courses may spend more time-on-task due to increased motivation. Other work shows potential for research-based instructional strategies to alter students' use of study time, encouraging deeper learning activities especially if course assessments emphasize this goal. Moreover, learning gains in courses using research-based instructional strategies cannot be explained solely by increased time-on-task. Regarding student response, large-scale works suggest that student response to interactive classrooms is largely positive. Each of these results would benefit from a new collection of studies including large-scale time use surveys, detailed time diary studies, and student opinion to better address questions of how specific research-based instructional strategies affect student study approach and student response.

Theme V. How will faculty find the time to adopt and maintain research-based pedagogy?

> '*I accept the value of interactive engagement techniques, but I do not have the additional time required to implement these approaches.*'
>
> '*I tried to use clickers, but I didn't see any improvements, so I returned to traditional lecturing.*'
>
> '*There are too many choices of teaching innovations; I don't know which to choose.*'

[10] Though students resist these devices being used solely for attendance monitoring.

In a survey of 722 physics faculty, Dancy and Henderson found that 'time' is the biggest impediment to implementing reforms (Dancy and Henderson 2010). In the free-response field to the question, 'What prevents you from using more of these [research-based] strategies?' 55.1% of faculty described a lack of time. This likely has two main contributing factors: the time to research, select, and adapt the pedagogies to be implemented, and the time required to maintain the strategy during the school year. However, Dancy and Henderson 'hesitate to conclude that with more time, instructors would implement more research-based strategies' and point out that often researchers invest considerable time in preparing detailed lecture notes and presentations. A challenge to the adoption of research-based instructional strategies is that they are often viewed as something to be added on top of an instructor's considerable workload (e.g., instead of preparing only one lecture, one needs to prepare several depending on the in-class feedback from students) rather than a replacement for some of the approaches currently in use. We look to the literature to better understand the most appropriate uses of the limited time available to an instructor, and the current feedback mechanisms (or lack of them) that aid instructors to improve their time management practices.

With limited time, an instructor needs to determine what activities and preparation are important for good teaching and what are less essential. From our experience, a typical instructor's high priorities will include picking a textbook, developing a detailed syllabus of content to be covered, creating and reviewing lecture notes, generating problem set questions with solutions, preparing lecture materials and demonstrations, and testing students' mastery of the material. Activities that often fall lower on the priority list include researching and implementing novel teaching techniques, adjusting instruction based on regular feedback from students, and comparing learning gains to other classes. Examples of activities that often end up at the very bottom of the priority list include seeking formalized teaching training, adjusting instruction techniques based on feedback from expert teachers, and measuring concept retention for students in subsequent courses.

A standard compliment for a well-prepared instructor illustrates many instructors' current time priorities: He/she did not need to look at their notes even once through the entire lecture! Masterful performance is essential for entertainment, but its effectiveness in learning is not clear. One perspective is that if a performance is too smooth, students are reduced to spectators with reduced critical thinking and therefore limited learning. There is no doubt that instructor preparation is essential to effective learning; the real question is, 'What preparation?' Which activities achieve the best learning outcomes for a given amount of time invested? One of the pressures at play is that often activities promoted by research-based pedagogies appear at odds with what a student, who admires the charismatic lecturer, might consider good teaching. For example, many research-based instructional strategies require that the instructor shifts the 'spotlight' from themselves to the students through Peer Instruction, hands-on self-discovery activities, or other student-centered activities. A student commented to one of us that they thought such activities were easier for the instructor because 'you do not have to teach as much'. This underscores the importance of using validated student learning assessments to overcome traditional misconceptions (e.g., students learn the most when an instructor clearly describes the concept) and provide accurate feedback on the effectiveness of a teaching strategy.

We make a case then that most post-secondary instructors have not had appropriate training (either before commencing their career or through continuing professional development) to determine the best set of time priorities for optimal student learning. The lack of a culture of teaching scholarship among practitioners aggravates the situation. Instructors are left to set priorities based on what they observed as a student. This would be similar to someone claiming to be an expert in planning a medical doctor's schedule of activities based on his observations as a patient. It is unreasonable for a researcher to suspend their research activities to undergo instructional development (and it is not clear whether many institutions offer appropriate post-secondary instructor training). While universities often have an 'Instructional Development Center', the role of such teaching centers is viewed (by some) to be limited because their mandate usually requires them to serve instructors from all disciplines. The concern becomes: How can someone who knows very little about physics help me to teach physics?

In setting time priorities, it is useful to consider another area of responsibility for most instructors: research. A logical approach to the problem of teaching time management is to apply successful methods taken by an instructor in their pursuit of research goals. The researcher's top priorities include undergoing high quality training and mentorship, developing a good understanding of the literature, satisfying the demands of external reviewers, having results verified by external experts, and training highly qualified personnel. This list stands in stark contrast to the instructor priority list described above. This is an important disconnect: the priority items for a researcher tend to end up at the bottom of the list for an instructor. A case could be made that because physics teaching and physics research are completely different activities, their priority lists should be different. However, in theme I, we demonstrated that effective teaching has considerable scholarship supporting its development and application. We see that it is beneficial to consider how a teacher can achieve better learning outcomes using a researcher's perspective on setting time priorities. In particular, a researcher's time management is strongly driven by very specific and timely feedback (reviewer comments, successful grant applications); this kind of feedback is essential and often lacking for the teacher.

Effective, timely, and constructive feedback about instructional successes and failures is essential for improved teaching, especially considering faculty's common lack of formal training. Teaching is typically evaluated by end-of-term student surveys. Because such results can be easily quantified, standardized, and collected, they have become the *de facto* standard teacher evaluation tool relevant for promotion and merit increases. Seldin (1998) states approximately

90% of American colleges report using them, even though there are studies that specifically question the validity of surveys (Kulik 2001). With a lack of confidence in their validity, their use to faculty as feedback is limited. Common complaints include, 'They are just a popularity contests' and 'If you want good reviews, just grade easily.' For interactive engagement, especially, there is a general feeling that student surveys can provide negative feedback. For example, Melissa Franklin (Harvard University) was quoted by the Chronicle of Higher Education (19 February 2012) to say 'The average score on a student evaluation of a flipped course is about half what the same professor gets when using the traditional lecture.' As discussed in theme IV, this remark is at odds with studies that show the majority of students prefer interactive engagement approaches (Gasiewski *et al* 2011). In addition, students who fully adapt to the interactive engagement approach tend to become strong proponents. Clearly, we require further (preferably) nationwide studies of student survey responses relative to specific active learning techniques and their implementations. Nonetheless, for a specific instructor aiming to improve his time management, end-of-term student surveys are of extremely limited value.

Another limiting factor in the effectiveness of current teaching feedback mechanisms is their (limited) role in reward structures. In a study of more than 4000 faculty at four-year colleges and universities, Fairweather found 'that actual rewards in terms of pay, tenure, and promotion are based almost exclusively on research productivity across all institutional types' (Fairweather 1993). This study includes institutions focused on undergraduate education. We recognize a second disconnect in the overall effort to improve student learning: even at institutions that have a strong teaching mandate, teaching performance is not the primary metric used for advancement. At root, we attribute this disconnect to a basic mistrust of standard evaluation techniques which rely heavily on 'client' reporting (student surveys, anonymous student testimonials/critiques) when we, as instructors, doubt many of our clients' ability to assess. Such feedback mechanisms are not well aligned to better education if student objectives are to 'get an A' instead of 'be challenged intellectually'. It is no surprise then that at the committee level, standard teaching metrics are mistrusted. The surprising result is that, even in the face of institutional indifference, faculty devote considerable time to teaching and the betterment of their students. Even while investing more time in research, faculty devote constant or increased time to teaching (Milem *et al* 2000), most likely because of intrinsic motivation. Nonetheless, a better alignment between reward structures and desired outcomes would likely increase instructor motivation to improve learning outcomes.

PER continues to explore new avenues for feedback to assist professors in prioritizing their teaching time. Peer review is the primary feedback mechanism for improving scholarly activity. From best practices of peer review, we explore what models might provide better feedback for the scholarly activity of teaching. Evans *et al* find that both review training and time spent on the review correlates with the quality of the review (Evans *et al* 1993), suggesting that teaching reviewers would require sufficient resources (particularly time) to be trained in evaluating teaching. A clear challenge to implementing peer review to teaching is providing reviewers with sufficient information about the course implementation and learning environment. The use of 'expert visits' for just one or two lectures is maligned as 'drive-by' observations. Scriven goes as far as stating that using such visits to evaluate teaching 'is not just incorrect, it is a disgrace' because of a variety of factors: the effect of the visit on the teaching, the limited observing time, expert biases, and the many other components of teaching that are not observed (Scriven 1981). Some institutions attempt to increase the quality of peer review by the use of a detailed teaching portfolio (Bernstein *et al* 2006). These dossiers have a role in job and tenure review files, but limited application as a method of providing timely formative assessment to a teacher as part of their regular professional development. Further effort is needed for PER to develop formative evaluation techniques for providing timely feedback to instructors.

When faculty do make use of feedback to connect an unsatisfactory level of learning in their classroom to their teaching approach, their teaching can improve greatly. However, effective time use is important, as all practitioners need not be creators. Faculty members can exploit current resources to achieve gains without the need to 'reinvent the wheel'. Again a parallel can be made to the researcher's perspective through the popular axiom 'Why waste an afternoon in the library, when I can spend six months in the lab figuring something out?' There is no need for the researcher to repeat all the pitfalls that have already been overcome. A researcher builds on prior contributions (such as those listed in the PER User's Guide (www.compadre.org/perug/) or previous redesigns by fellow faculty who already have incorporated active learning. Taking this approach in teaching would allow instructors to optimize effective teaching strategies without wasting time repeating common mistakes. This approach is particularly important for teaching because a misguided practice can have serious negative outcomes for students and can take months to be corrected.

A way forward is suggested by the positive results achieved by Science Education Specialists through the University of British Columbia Carl Wieman Science Education Initiative who served as embedded education experts working within departments (Wieman *et al* 2013). These experts (in both the specific discipline as well as in discipline-based education research) provide instructors expert advice on best practices for course development and detailed formative feedback based on classroom observations as well as interviews with students. 69 out of 70 instructors who implemented research-based instructional strategies through the Initiative continued using these strategies extensively, even well after their interaction with a Science Education Specialist had ended. The scope of this multi-million dollar initiative to transform science education is challenging to replicate and as the authors note, further research is required to determine 'what support is the most important of everything that is provided by the CWSEI, and what is the most cost-effective way to provide that support.'

Effective time investment is necessary for optimized active learning courses. Accessing the literature on research-based instructional strategies can allow efficient adoption of proven techniques while avoiding the pitfalls that the developers experienced and overcame. One issue that is not well documented in PER is the amount of time that is required to implement and maintain an instructional strategy. Studies usually carefully measure the effectiveness of the pedagogy in terms of conceptual understanding or other learning outcomes, but it is often difficult to make an estimate of the 'rate on return' for time invested by the faculty member. This can be done relatively easily, even if not rigorously. For example, Deslauriers *et al* state that the active engagement strategies they evaluated initially took twenty hours of preparation for one hour of class, but this dropped to ten hours by the third lecture (Deslauriers *et al* 2011). They hypothesize that an experienced teacher could achieve similar results in five hours, with dramatic reductions in future years due to the reuse of materials (about two hours of for each hour of lecture). Just in Time Teaching is another research-based instructional strategy that instructors often identify as being too time consuming to maintain (Novak *et al* 1999). The standard implementation requires students to complete reading assignments before lectures; the instructor reviews this feedback to set that day's topics. In a large enrolment course, an instructor could not reasonably review 200–300 responses before every class. However, in the same way that an experimentalist would measure only a random subset of a large data set to determine some likely 'average value', the instructor can sample a subset of the reading responses. PER does not define a 'good' sampling rate (taking into consideration not only accuracy of the assessment, but also the effect of limited sampling on student attitudes and motivation to do subsequent reading assignments). It is in cases like this that physics education literature fails the interested faculty member. Physics education researchers should respond to practical concerns that are essential to successful adoption of research-based instructional strategies.

A final issue of instructor time is determining the appropriate amount of time to devote to a course. It is unreasonable for instructors to claim that they do not have enough time if they have not considered in a quantitative sense, how many hours per week they should devote to teaching. This number is very dependent on an instructor's individual situation, but a rough baseline can be estimated. A typical workload at a research institution would require about 40% of the researcher's time be devoted to teaching (with the remainder for research and service work). A quick calculation (assuming 40 h work weeks and 48 weeks/year) suggests that a minimum of 768 h should be devoted to teaching. A recent study suggested that researchers work considerably more time than 40 h/week, but in this analysis we restrict ourselves to the bare minimum. If the instructor is to provide three term courses of instruction (each with a 14-week term), our simple calculation suggests they should be devoting over 18 h/week to each course. Assuming three hours of lecture time, we are left with 15 h/week. Some of these hours need to be spent before the term begins for preparation, and during the exam period for marking and course administration. Nonetheless such a simple analysis can be done by every instructor for their particular situation to allow them to adjust their time priorities accordingly. It is possible that the perceived lack of time for improved teaching reported by many instructors is due to other responsibilities encroaching into their teaching time.

In summary, effective time management is essential to improved teaching. Instructors need to open feedback channels to provide themselves with timely feedback on the progress (or lack thereof) of student learning. By taking a research approach to improving teaching, instructors can optimize their time to efficiently achieve improved learning gains. Physics education researchers must also work to answer outstanding questions, including estimates of the time required for particular pedagogy implementations.

2. Conclusion

Physics education research has potential to benefit students and society by successfully training the skilled science professionals that our economy needs. Furthermore, in our changing academic landscape where the professor is no longer the main disseminator of information, it is in the best interest of physics instructors to adopt research-based teaching strategies and provide the greatest possible learning experiences for their students. However, despite the benefits and importance of applying research-based pedagogy in physics classrooms, many professors have not made use of the physics education resources available to them. In this report, we have discussed five common barriers to the adoption of research-based instructional strategies.

First, the need for scientific rigor in evaluating and improving physics instruction was discussed under theme I. Physics education research generates the most generalizable conclusions through meta-analyses on repeated experiments with validated measuring instruments. Anecdotes and recollections of one's own experiences as a learner can inspire a professor's teaching, but they can also dangerously bias it if these anecdotes are not used in conjunction with rigorous physics education research. Physics education research has measured the learning gains of students in traditional courses compared to active learning courses and found strong evidence for the success of interactive pedagogies, as discussed under theme II. Research-based instructional strategies have been shown to be significantly more successful than the traditional lecture for improving students' conceptual understanding, engagement, and retention in physics programs. Furthermore, these important gains need not come at the cost of inhibiting students' skills—in fact, these gains may improve problem-solving skills, as described under theme III.

Time is often the bottom line for both professors and students. As discussed under themes IV and V, careful time allocation is crucial for successful implementations of research-based instructional strategies. Both student and professor time can be wasted through unsuccessful attempts to transfer knowledge: for example, a professor could spend hours crafting elegant lecture notes that students merely copy into their notebook with limited understanding or critical thought. Instead, physics education research seeks to provide both instructors and students with strategies to maximize successful and lasting learning experiences within their limited time.

It is encouraging to note that the use of research-based instructional strategies allows students to achieve higher learning gains in equivalent student time by altering the ways students invest their time, as described in theme IV. Nonetheless, students could be resistant to such changes. Therefore, we highlight the importance of building student trust and buy-in when adopting a new teaching strategy. Active learning pedagogies can increase student motivation and enjoyment in the course when the students fully understand the utility of their work.

Just as research-based pedagogies can change how students spend their time, instructors can make use of physics education research to prioritize their time, as discussed in theme V. Physics education literature can guide instructors in setting time priorities so that instructors do not need to waste time 'reinventing the wheel'. By taking a research approach to their teaching—consulting the literature, undergoing professional development opportunities, and verifying their results using standard validated assessments—professors can make use of instructional strategies that have already been tested and improved. In order for professors to select appropriate pedagogies that work within the constraints of their allotted teaching time, physics education researchers should measure and report the time required to implement specific research-based teaching strategies. It would be prudent for university administrations to re-evaluate their emphasis on rewarding research productivity over teaching accomplishments, so that administrations can better support faculty in their teaching professional development.

As physics researchers, it is logical that we should take a research approach to improving our teaching. One place to start is through opening at least one additional feedback channel (e.g., ask a few conceptual questions during each lecture and obtain responses from *every* student). As this additional feedback provides a clearer picture of how and what students are learning (and not learning) in current classrooms, instructors can make use of the literature and professional development opportunities to master research-based pedagogies. Resources such as the Physics Education Research User's Guide and communities such as peerinstruction.net can support professors as they optimize research-based instructional strategies for their specific classrooms. Physics education research has the potential to yield significant benefits for students, professors, and society. With the considerable shift in the educational landscape catalyzed by changes in both technological innovations and budget pressures, we need to adopt active engagement teaching strategies to avoid obsolescence and best serve our students, one student at a time.

Acknowledgments

Several people contributed to this manuscript. EM led the research team. JF conceived of the structure and organized motivating questions for each section. JF, KM, JD, LT and EM developed the initial outline. JF, AT, KM, JD and LT contributed equally to the first draft of the body of the manuscript. AT compiled contributions and wrote the first draft of the conclusion. All authors subsequently took part in the revision process and approved the final copy of the manuscript.

References

Adams W K, Perkins K K, Podolefsky N S, Dubson M, Finkelstein N D and Wieman C E 2006 New instrument for measuring student beliefs about physics and learning physics: the Colorado Learning Attitudes about Science Survey *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **2** 010101

Anders Ericsson K, Krampe R Th and Tesch-Romer C 1993 The role of deliberate practice in the acquisition of expert performance *Psychol. Rev.* **100** 363–406

Bernard R, Abrami P, Borokhovski E, Lou Y, Wozney L, Wade A, Fiset M, Wallet P A and Huang B 2004 How does distance education compare with classroom instruction? A meta-analysis of the empirical literature *Rev. Educ. Res.* **74** 379–439

Bernstein D, Burnett A, Goodburn A and Savory P 2006 *Making Teaching and Learning Visible: Course Portfolios and the Peer Review of Teaching* (Bolton, ME: Anker)

Bing T and Redish E 2009 Analysing problem solving using math in physics: epistemological framing via warrants *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **5** 020108

Bol L, Warkentin R W, Nunnery J A and O'Connell A A 1999 College students' study activities and their relationship to study context, reference course, and achievement *Coll. Stud. J.* **33** 608

Bolger N, Davis A and Rafaeli E 2003 Diary methods: capturing life as it is lived *Annu. Rev. Psychol.* **54** 579–616

Boyle J T and Nicol D J 2003 Using classroom communication systems to support interaction and discussion in large class settings *Assoc. Learn. Technol. J.* **11** 43–57

Bressoud D 2011 *The Calculus I Instructor, Launchings* (June) www.maa.org/columns/launchings/launchings_06_11.html

Brewe E, Kramer L and O'Brien G 2009 Modeling instruction: positive attitudinal shifts in introductory physics measured with CLASS *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **5** 013102

Brewe E, Sawtelle V, Kramer L H , O'Brien G E, Rodriguez I and Pamelá P 2010 Toward equity through participation in modeling instruction in introductory University Physics *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **6** 010106

Brewe E, Kramer L and O'Brien G 2009 Modeling instruction: positive attitudinal shifts in introductory physics measured with CLASS *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **5** 013102

Brewe E, Kramer L H and Sawtelle V 2012 Investigating student communities with network analysis of interactions in a physics learning center *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **8** 010101

Caballero M D 2011 Evaluating and extending a novel course reform of introductory mechanics *PhD Thesis* Georgia Institute of Technology

Caldwell J E 2007 Clickers in the large classroom: current research and best-practice tips *CBE Life Sci. Educ.* **6** 9–20

Chasteen S V, Pepper R E, Pollock S J and Perkins K K 2012 But does it last? Sustaining a research-based curriculum in upper-division electricity & magnetism *2010 Physics Education Research Conf. (Portland, OR) AIP Conf. Proc.* **1413** 139 www.Compadre.Org/portal/items/detail.Cfm

Chi M T H, Glaser R and Rees E 1982 Expertise in problem solving *Advances in the Psychology of Human Intelligence* vol 1 ed R J Sternberg (Hillsdale, NJ: Erlbaum)

Cooper H, Hedges L V and Valentine J C 2009 *Handbook of Research Synthesis and Meta-Analysis* (New York: Russell Sage Foundation)

Cornelius-White J 2007 Learner-centered teacher–students relationships are effective: a meta-analysis *Rev. Educ. Res.* **77** 113–43
Crouch C H and Mazur E 2001 Peer Instruction: ten years of experience and results *Am. J. Phys.* **69** 970
Dancy M and Henderson C 2010 Pedagogical practices and instructional change of physics faculty *Am. J. Phys.* **78** 1056
Denzin N and Lincoln Y 2005 *The Sage Handbook of Qualitative Research* 3rd edn (Thousand Oaks, CA: Sage Publications Inc.)
Derting T, Williams K S, Momsen J L and Henkel T P 2011 Education research: set a high bar *Science* **333** 1220–1
Deslauriers L, Schelew E and Wieman C 2011 Improved learning in a large-enrollment physics class *Science* **332** 862–4
Deslauriers L and Wieman C 2011 Learning and retention of quantum concepts with different teaching methods *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **7** 010101
Ding L, Chabay R, Beichner R and Sherwood B 2006 Evaluating an electricity and magnetism assessment tool: brief electricity and magnetism assessment *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **2** 010105
Elby A 1999 Another reason that physics students learn by rote *Am. J. Phys.* **67** S52–7
Evans A, Mcnutt R, Fletcher S and Fletcher R 1993 The characteristics of peer reviewers who produce good-quality reviews *J. Gen. Intern. Med.* **8** 422–8
Fairweather J S 1993 Faculty reward structures: toward institutional and professional homogenization *Res. Higher Educ.* **34** 603–23
Finkelstein N 2013 Affect not as an afterthought *AAPT National Meeting (Portland, OR, 17 July)* www.compadre.org/per/perc/2013/files/PERC2013_InvitedTalk_Finkelstein.pdf
Gasiewski J A, Eagan M K, Garcia G A, Hurtado S and Chang M J 2011 From gatekeeping to engagement: a multicontextual, mixed method study of student academic engagement in introductory STEM courses *Res. Higher Educ.* **53** 229–61
Hake R R 1998 Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses *Am. J. Physics* **66** 64–74
Halloun I A and Hestenes D 1985 The initial knowledge state of college physics students *Am. J. Phys.* **53** 1043–55
Hazari Z, Sonnert G, Sadler P M and Shanahan M-C 2010 Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: a gender study *J. Res. Sci. Teach.* **47** 978–1003
Hazari Z, Tai R H and Sadler P M 2007 Gender differences in introductory university physics performance: the Influence of high school physics preparation and affective factors *Sci. Educ.* **91** 847–76
Heckler A F and Sayre E C 2010 What happens between pre- and post-tests: multiple measurements of student understanding during an introductory physics course *Am. J. Phys.* **78** 768–77
Hegde B and Meera B 2012 How do they solve it? An insight into the learner's approach to the mechanism of physics problems solving *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **8** 010109
Heller J I and Reif F 1984 Prescribing effective human problem-solving processes: problem description in physics *Cogn. Instr.* **1** 177–216
Heller P and Hollabaugh M 1992 Teaching problem solving through cooperative grouping: II. Designing problems and structuring groups *Am. J. Phys.* **60** 367–44
Heller P, Keith R and Anderson S 1992 Teaching problem solving through cooperative grouping: I. Group versus individual problem solving *Am. J. Phys.* **60** 627–36
Henderson C and Dancy M 2009 Impact of physics education research on the teaching of introductory quantitative physics in the United States *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **5** 020107
Henderson C, Dancy M and Niewiadomska-Bugaj M 2012 Use of research-based instructional strategies in introductory physics: where do faculty leave the innovation-decision process? *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **8** 020104
Hertado S, Eagan K, Pryor J, Whang H and Tran S 2012 Undergraduate Teaching Faculty: The 2010–2011 HERI Faculty Survey *Higher Education Research Institute at UCLA*
Hessling R M, Traxel N M and Schmidt T J 2004 *The SAGE Encyclopedia of Social Science Research Methods* ed M S Lewis-Beck *et al* (Thousand Oaks, CA: Sage Publications, Inc.)
Hestenes D, Wells M and Swackhamer G 1992 Force concept inventory *Phys. Teach.* **30** 141–58
Heuvelen A V 1991a Learning to think like a physicist: a review of research-based instructional strategies *Am. J. Phys.* **59** 891–7
Heuvelen A V 1991b Overview, case study physics *Am. J. Phys.* **59** 898–907
Hoellwarth C, Moelter M J and Knight R D 2005 A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms *Am. J. Phys.* **73** 459–62
Jones L R, Miller A G and Watts J F 2000 Conceptual teaching and quantitative problem solving: friends or foes? *J. Cooperation Collaboration Coll. Teach.* **10** 109–16
Judson E and Sawada D 2002 Learning from past and present: electronic response systems in college lecture halls *J. Comput. Math. Sci. Teach.* **21** 167–81
Kember D, Jamieson Q W, Pomfret M and Wong E T T 1995 Learning approaches, study time and academic performance *Higher Educ.* **29** 329–43
Knight J K and Wood W B 2005 Teaching more by lecturing less *Life Sci. Educ.* **4** 298
Kulik J A 2001 Student ratings: validity, utility, and controversy *New Directions Inst. Res.* **2001** 9–25
Larkin J, McDermott J, Simon D P and Simon H A 1980 Expert and novice performance in solving physics problems *Science* **208** 1335–42
Larkin J and Reif F 1979 Understanding and teaching problem-solving in physics *Eur. J. Sci. Educ.* **1** 191–203
Lasry N, Charles E, Whittaker C and Lautman M 2009 When talking is better than staying quiet *AIP Conf. Proc.* **1179** 181–4
Lasry N, Mazur E and Watkins J 2008 Peer Instruction: from Harvard to the two-year college *Am. J. Phys.* **76** 1066
Leonard S, Stanne M E and Donovan S 1999 Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis *Rev. Educ. Res.* **69** 21–51
Leonard W J, Dufresne R J and Mestre J P 1996 Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in solving problems *Am. J. Phys.* **64** 1495–503
Maloney D, O'Kuma T, Hieggelke C and Heuvelen A V 2001 Surveying students' conceptual knowledge of electricity and magnetism *Am. J. Phys.* **69** S12–23
Marx C and Cummings K 2010 Development of a survey instrument to gauge students' problem-solving abilities *AIP Conf. Proc.* **1289** 221–4
Mauk H V and Hingley D 2005 Student understanding of induced current: using tutorials in introductory physics to teach electricity and magnetism *Am. J. Phys.* **73** 1164–71
McDermott L C and Shaffer P S 2001 *Tutorials in Introductory Physics* (Englewood Cliffs, NJ: Prentice-Hall)
Meltzer D E and Manivannan K 2002 Transforming the lecture-hall environment: the fully interactive physics lecture *Am. J. Phys.* **70** 639 (Abstract)
Mestre J P, Dufresne R J, Gerace W J, Hardiman P T and Touger J S 1993 Promoting skilled problem-solving behavior among beginning physics students *J. Res. Sci. Teach.* **30** 303–17
Milem J F, Berger J B and Dey E L 2000 Faculty time allocation: a study of change over twenty years *J. Higher Educ.* **71** 454–75

Mokhtari K, Reichard C A and Gardner A 2009 The impact of internet and television use on the reading habits and practices of college students *J. Adolescent Adult Literacy* **52** 609–19

Nonis S A and Hudson G I 2010 Performance of college students: impact of study time and study habits *J. Educ. Bus.* **85** 229–38

Novak G, Patterson E, Gavrin A and Christian W 1999 *Just-in-Time Teaching: Blending Active Learning with Web Technology* (Englewood Cliffs, NJ: Prentice-Hall)

Physics Education Research User's Guide. see www.compadre.org/perug/

Priest A G and Lindsay R O 1992 New light on novice—expert differences in physics problem solving *Br. J. Psychol.* **83** 389–405

Prince M 2004 Does active learning work? A review of the research *J. Eng. Educ.* **93** 223–32

Redish E F, Saul J and Steinberg R 1997 On the effectiveness of active-engagement microcomputer-based laboratories *Am. J. Phys.* **65** 45

Ruiz-Gallardo J-R, Castaño S, Gómez-Alday J J and Valdés A 2011 Assessing student workload in problem based learning: relationships among teaching method, student workload and achievement. a case study in natural sciences *Teach. Teach. Educ.* **27** 619–27

Sadaghiani H R 2011 Using multimedia learning modules in a hybrid-online course in electricity and magnetism *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **7** 010102

Sayre E C, Franklin S V, Dymek S, Clark J and Sun Y 2012 Learning, retention, and forgetting of Newton's third law throughout university physics *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **8** 010116

Schwartz D L and Bransford J D 1998 A time for telling *Cogn. Instr.* **16** 475–522

Scriven M 1981 Summative teacher evaluation *Handbook of Teacher Evaluation* ed J Millman (Beverly Hills, CA: Sage) p 251

Seldin P 1998 How colleges evaluate teaching: 1988 vs 1998 *Am. Assoc. Higher Educ. Bull.* **50** 3–7

Smith M K, Wood W B, Krauter K and Knight J K 2011 Combining Peer discussion with instructor explanation increases student learning from in-class concept questions *CBE Life Sci. Educ.* **10** 55–63

Smith M U (ed) 1991 *Toward a Unified Theory of Problem Solving: Views from the Content Domains* (Hillsdale, NJ: Erlbaum)

Springer L, Stanne M E and Donovan S S 1999 Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis *Rev. Educ. Res.* **69** 21–51

Stelzer T, Brookes D T, Gladding G and Mestre J P 2010 Impact of multimedia learning modules on an introductory course on electricity and magnetism *Am. J. Phys.* **78** 755

Stewart J, Stewart G and Taylor J 2012 Using time-on-task measurements to understand student performance in a physics class: a four-year study *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **8** 010114

Taconis R, Ferguson-Hessler M G M and Broekkamp H 2001 Teaching science problem solving: an overview of experimental work *J. Res. Sci. Teach.* **38** 442–68

Thacker B, Kim E, Trefz K and Lea S M 1994 Comparing problem solving performance of physics students in inquiry-based and traditional introductory physics courses *Am. J. Phys.* **62** 627–33

Thambiratnam D P, Ferreira L and Franz J M 1994 A Phenomenographic Approach to Understanding Student Learning: Some Case Studies from Civil Engineering http://eprints.qut.edu.au/7798/

Thornton R and Sokoloff D 1998 Assessing student learning of Newton's laws: the force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula *Am. J. Phys.* **66** 338

Torgerson C 2011 Education research: call for controls *Science* **333** 1220

Turpen C and Finkelstein N 2009 Not all interactive engagement is the same: variations in physics professors' implementation of Peer Instruction *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **5** 020101

Turpen C and Finkelstein N D 2010 The construction of different classroom norms during Peer Instruction: students perceive differences *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **6** 020123

Turpen C, Finkelstein N D and Pollock S J 2009 Towards understanding classroom culture: students' perceptions of tutorials *AIP Conf. Proc.* **1179** 285–88

Ward M and Sweller J 1990 Structuring effective worked examples *Cogn. Instr.* **7** 1–39

Watkins J and Mazur E 2013 Retaining students in science, technology, engineering, and mathematics (STEM) majors *J. Coll. Sci. Teach.* **42** 36–41

Wells M, Hestenes D and Swackhamer G 1995 A modelling method for high school physics instruction *Am. J. Phys.* **63** 606

Weinburgh M 1995 Gender differences in student attitudes toward science: a meta-analysis of the literature from 1970 to 1991 *J. Res. Sci. Teach.* **32** 387–98

Wieman C, Deslauriers L and Gilley B 2013 Use of research-based instructional strategies: how to avoid faculty quitting *Phys. Rev. Spec. Top.—Phys. Educ. Res.* **9** 023102

Wright D S and Williams C D 1986 A wise strategy for introductory physics *Phys. Teach.* **24** 211–6

Zuriff G E 2003 A method for measuring student study time and preliminary results *Coll. Stud. J.* **37** 72–8