Six Lessons From The Physics Education Reform Effort *†

Richard Hake <rrhake@earthlink.net>, <http://www.physics.indiana.edu/~hake>
Physics Department, Indiana University, Emeritus

In a 1998 meta-analysis I showed that “interactive engagement” (IE) courses could yield average normalized pre-to-posttest gains $g$ in conceptual understanding of Newtonian mechanics that were about two standard deviations greater than traditional (T) courses. Then in 2002 I wrote a paper based on my meta-analysis entitled “Lessons From the Physics Education Reform Effort.” There, among other things, I offered six lessons on “interactive engagement” that I had hoped might stimulate more effective high school and university education. Today five years latter, it may be worthwhile to review and update those lessons with an eye to the present status of education reform in physics and other disciplines.

Introduction

In a meta-analysis titled “Interactive-engagement vs traditional methods: A six thousand-student survey of mechanics test data for introductory physics courses” [Hake (1998a,b)], I showed that interactive engagement (IE) courses could yield average normalized pre-to-posttest gains $g$ in conceptual understanding of Newtonian mechanics that were about two standard deviations greater than traditional (T) courses. Here:

a. The average normalized gain $g$ is the average actual gain $[\%post - \%pre]$ divided by the maximum possible average gain $[100\% - \%pre]$, where the angle brackets indicate the class averages.

b. The conceptual tests of Newtonian Mechanics were either the Force Concept Inventory (FCI) [Hestenes et al. (1992)] or its precursor the Mechanics Diagnostic (MD) [Halloun & Hestenes (1985a,b)]; in both cases developed by disciplinary experts through arduous qualitative and quantitative research, and widely recognized as valid and consistently reliable.

c. IE courses were operationally defined as those designed at least in part to promote conceptual understanding through continual interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback through discussion with peers and/or instructors.

d. T courses were operationally defined as those reported by instructors to make little or no use of IE methods, relying primarily on passive-student lectures, recipe labs, and algorithmic problem exams.


†Partially supported by NSF Grant DUE/MDR-9253965.
In a later paper “Lessons from the physics education reform effort” [Hake (2002a)] I offered fourteen lessons, six on interactive engagement and eight on implementation, that I hoped might stimulate more effective high-school and university education. Today, five years latter, it may be worthwhile to review, update, and add to those lessons with an eye to the present status of education reform in physics and other disciplines. In the present paper I shall discuss only the six lessons on interactive engagement.

These lessons are derived from my own interpretation of the physics education reform movement as it has developed over the past few decades, and are, therefore, somewhat subjective and incomplete. They are meant to stimulate discussion rather than present any definitive final analysis.

**LESSON 1: The use of Interactive Engagement (IE) strategies can increase the effectiveness of conceptually difficult courses well beyond that obtained by traditional (T) methods.**

Education research in chemistry [Krause et al. (2004)]; engineering [Froyd et al. (2006), Evans et al. (2003)]; and introductory science education generally [Handelsman et al. (2004)], although neither as extensive nor as systematic as that in physics [McDermott and Redish (1999); Redish (1999); Thacker (2003); Heron & Meltzer (2005); Hake (1998a,b; 2002a,b; 2005a; 2006a,b; 2007a,b); Wieman & Perkins (2005); Wieman (2005)] is consistent with the latter in suggesting that, in conceptually difficult areas, Interactive Engagement (IE) methods are more effective than traditional T passive-student methods in enhancing students' understanding. Furthermore, there is some preliminary evidence that learning in IE physics courses is substantially retained 1 to 3 years after the courses have ended [Chabay (1997), Francis et al. (1998), Bernhard (2001)].

I see no reason to doubt that enhanced understanding and retention would result from greater use of IE methods in other science, and even non-science, areas, but substantive research on this issue is sorely needed – see e.g., “The Physics Education Reform Effort: A Possible Model for Higher Education?” [Hake (2005a)].

Pre/post testing in biology [Klymkowsky et al. (2003), Klymkowsky (2007)]; and mathematics [Epstein (2005)] is just getting started; while pre/post test results in astronomy (Brogt et al. (2007) and geoscience [Libarkin & Anderson (2005)], have not, at this early stage, shown clear-cut correlations between pre-to-posttest gain and pedagogical method, as has been shown in physics.
**LESSON 2: The use of IE and/or high-tech methods, by themselves, does not ensure superior student learning.**

The data shown in Fig. 1 of “Interactive-engagement vs traditional methods: A six thousand-student survey of mechanics test data for introductory physics courses” [Hake (1998a)], indicate that seven of the IE courses (717 students) achieved <g>'s close to those of the T courses. Five of those made extensive use of high-tech microcomputer-based labs [Thornton and Sokoloff (1990, 1998)]. Case histories of the seven low-<g> courses (Hake 1998b) suggest that implementation problems occurred.

It should be emphasized that, although high technology is, by itself, no panacea, it can be very advantageous when it promotes interactive engagement, as in:

(a) computerized classroom communication systems - see, e.g., Bruff (2007) and Hake (2007e).

(b) *properly implemented* microcomputer-based labs [Thornton and Sokoloff (1990)];

(c) computer-implemented tutorials [Reif and Scott (1999)];

(e) *Just-in-time teaching* [Novak et al. (1999), Gavrin (2001)].

**LESSON 3: High-quality standardized tests of the cognitive and affective impact of courses are essential to gauge the relative effectiveness of non-traditional educational methods.**

So great is the inertia of the educational establishment that three decades of physics education research [McDermott and Redish (1999)] demonstrating the futility of the passive-student lecture in introductory courses was ignored until Halloun and Hestenes (1985a,b) devised the *Mechanics Diagnostic* (MD) test of conceptual understanding of Newtonian mechanics.

Among many other virtues, the MD and the subsequent *Force Concept Inventory* (FCI) (Hestenes et al. 1992) tests have two major advantages: (a) the multiple-choice format facilitates relatively easy administration of the tests to thousands of students; (b) the questions probe for a conceptual understanding of the basic concepts of Newtonian mechanics in a way that is understandable to the novice who has never taken a physics course, yet at the same time are rigorous enough for the initiate.

Thus the questions can be given as an introductory course pretest in pre/post tests to directly determine course-induced gain in conceptual understanding. In my opinion such *direct* gain measurements of higher-order student learning are far superior to the *indirect* (and therefore in my view problematic) gauges have been developed; e.g., Reformed Teaching Observation Protocol (RTOP), National Survey Of Student Engagement (NSSE), Student Assessment of Learning Gains (SALG), and Knowledge Surveys (KS’s) [Nuhfer & Knipp( 2003)]. For a discussion and references for all but the last see Hake (2005b.)

BUT WAIT!
1. Can multiple choice tests gauge higher level cognitive outcomes such as the conceptual understanding of Newtonian mechanics? Wilson & Bertenthal (2005) think so, writing (p. 94):

Performance assessment is an approach that offers great potential for assessing complex thinking and learning abilities, but multiple choice items also have their strengths. For example, although many people recognize that multiple-choice items are an efficient and effective way of determining how well students have acquired basic content knowledge, many do not recognize that they can also be used to measure complex cognitive processes. For example, the Force Concept Inventory . . . [Hestenes, Wells, & Swackhamer, 1992] . . . is an assessment that uses multiple-choice items to tap into higher-level cognitive processes.

2. Considering the canonical arguments regarding the invalidity of pre/post testing evidence, should not all pre-to-post test gains cited above be viewed with grave suspicion? The dour appraisal of pre/post testing by Cronbach & Furby (1970) has echoed down though the literature to present day texts on assessment such as that by Suskie (2004)]. In my opinion, such pre/post paranoia and its attendant rejection of pre/post testing in evaluation, as used so successfully in physics education reform [McDermott and Redish (1999); Redish (1999); Thacker (2003); Heron & Meltzer (2005); Hake (1998a,b; 2002; 2005a; 2006a,b; 2007a,b); Wieman & Perkins (2005), Wieman (2005)] is one reason for the glacial progress of educational research [Lagemann (2000)] and reform [Bok (2005)].

Fortunately formative pre/post testing is gradually gaining a foothold in undergraduate astronomy, biology, chemistry, economics, geoscience, and engineering, in addition to physics. For references see Hake (2004, 2007c, 2007d).

Regarding tests of affective impact:
(a) administration of the Maryland Physics EXpectations (MPEX) survey to 1500 students in introductory calculus-based physics courses in six colleges and universities showed "a large gap between the expectations of experts and novices and . . . a tendency for student expectations to deteriorate rather than improve as a result of introductory calculus-based physics" [Redish et al. (1998)]. Here the term "expectations" is used to mean a combination of students' epistemological beliefs about learning and understanding physics and students' expectations about their physics course [Elby (1999), Elby (2001)] has recently conducted classes designed to help students develop more sophisticated beliefs about knowledge and learning as measured by MPEX.

(b) The Arizona State University "Views About Sciences Survey" (VASS) [Halloun and Hestenes (1998)], (available for physics, chemistry, biology, and mathematics at <http://modeling.la.asu.edu/R&E/Research.html>) indicates that students have views about physics that (i) often diverge from physicists' views; (ii) can be grouped into four distinct profiles: expert, high transitional, low transitional, and folk; (iii) are similar in college and high school; and (iv) correlate significantly with normalized gain g on the FCI. It may well be that students' attitudes and understanding of science and education are irreversibly imprinted in the early years [but see Elby (2001)]. If so, corrective measures await a badly needed shift of K–12 education away from rote memorization and drill (often encouraged by state-mandated standardized tests) to the enhancement of understanding and critical thinking [Mahajan and Hake (2000)].
(c) The “Colorado Learning Attitudes about Science Survey” [Adams et al. (2006), according to the abstract:]

... serves as the foundation for an extensive study of how student beliefs impact and are impacted by their educational experiences. For example, this survey measures the following: that most teaching practices cause substantial drops in student scores; that a student’s likelihood of becoming a physics major correlates with their “Personal Interest” score; and that, for a majority of student populations, women’s scores in some categories, including “Personal Interest” and “Real World Connections,” are significantly different from men’s scores.

**LESSON 4: Education Research and Development (R&D) by disciplinary experts (DE’s), and of the same quality and nature as traditional science/engineering R&D, is needed to develop potentially effective educational methods within each discipline. But the DE’s should take advantage of the insights of (a) DE’s doing education R&D in other disciplines, (b) cognitive scientists, (c) faculty and graduates of education schools, and (d) classroom teachers.**

Redish (1999) has marshaled the arguments for the involvement of physicists in physics departments, not just faculty of education schools, in physics education research. Similar arguments may apply to other disciplines. For physics, Redish gave these arguments:

(a) physicists have good access to physics courses and students on which to test new curricula;

(b) physicists and their departments directly benefit from physics education research;

(c) education schools have limited funds for disciplinary education research; and

(d) understanding what's going on in physics classes requires deep rethinking of physics and the cognitive psychology of understanding physics.

One might add that the researchers themselves must be excellent physics teachers with both content and "pedagogical content" knowledge [Shulman (1986, 1987)] of a depth unlikely to be found among non-physicists.

The education of disciplinary experts in education research requires PhD programs at least as rigorous as those for experts in traditional research. The programs should include, in addition to the standard disciplinary graduate courses, some exposure to: the history and philosophy of education, computer science, statistics, political science, social science, economics, engineering, and, most importantly, cognitive science (i.e., philosophy, psychology, artificial intelligence, linguistics, anthropology, and neuroscience). In my opinion, all scientific disciplines should consider offering PhD programs in education research.
As far as I know, physics leads the way in preparing future educational researchers and in researching undergraduate student learning – see e.g. Stockstad (2001). For links to over 50 U.S. Physics Education Research (PER) groups (many of them offering Ph.D.’s), over 200 PER papers published in the American Journal of Physics since 1972, and tests of cognitive and affective conditions see, respectively: Meltzer (2002), Meltzer (2005), and NCSU 2005. The very active PER discussion list PhysLrnR <http://listserv.boisestate.edu/archives/physlrnr.html> logged over 1100 posts in 2006. To access the archives of PhysLnR one needs to subscribe, but that takes only a few minutes by clicking on
<http://listserv.boisestate.edu/archives/physlrnr.html> and then clicking on "Join or leave the list (or change settings)." If you're busy, then subscribe using the "NOMAIL" option under "Miscellaneous." Then, as a subscriber, you may access the archives and/or post messages at any time, while receiving NO MAIL from the list!

**LESSON 5: The development of effective educational methods within each discipline requires a redesign process of continuous long-term classroom use, feedback, assessment, research analysis, and revision.**

Wilson and Daviss (1994) suggest that the "redesign process," used so successfully to advance technology in aviation, railroads, automobiles, and computers can be adapted to K–12 education reform through "System Redesign Schools." Redesign processes in the reform of introductory undergraduate physics education have been undertaken and described by McDermott (1991) and by Hake (1998a). In my opinion, "redesign" at both the K–12 and undergraduate levels can be greatly assisted by the promising "Scholarship of Teaching & Learning" movement - see e.g., Carnegie Academy (2007) and IJ-SOTL (2007).

**LESSON 6: Although non-traditional IE methods appear to be much more effective than T methods, there is need for more research to develop better strategies for enhancing student learning.**

On a test as elemental as the FCI, it would seem that reasonably effective courses should yield normalized gains $g$'s above 0.8, but thus far very few above 0.7 have, to my knowledge, been reported. This and the poor showing on the pre/post MPEX test of student understanding of the nature of science and education (Redish et al. 1998) indicate that more work needs to be done to improve IE methods. It would seem that understanding of science might be enhanced by:

(a) students' apprenticeship research experiences (Collins et al. 1989, Brown et al. 1989);

(b) epistemologically oriented teachers, materials, and class activities [Elby (1999, 2001); May & Etkina (2002); Hammer & Elby (2003), Lising & Elby (2005)];

(c) enrollment in courses featuring interactive engagement among students and disciplinary experts from different fields, all in the same classroom at the same time [Benbasat and Gass (2001)];

(d) further investigation of the connection between scientific reasoning ability and normalized gain on conceptual tests [see e.g., Coletta & Phillips (2005); Coletta, Phillips, & Steinert (2007a,b)];
(e) better communication between educational researchers and cognitive scientists – see e.g. “Cognitive Science and Physics Education Research: What we’ve got here is a failure to communicate ” [Hake (2007f)].

(f) multifaceted assessments – see e.g. Etkina et al. (2006) - to gauge the effectiveness of introductory courses in promoting students’ capacities [Hake (1998b)] in areas other than conceptual understanding: e.g., students’:

(a) satisfaction with and interest in physics;
(b) understanding of the nature, methods, and limitations of science;
(c) understanding of the processes of scientific inquiry such as experimental design, control of variables dimensional analysis, order-of-magnitude estimation, thought experiments, hypothetical reasoning, graphing, and error analysis;
(d) ability to articulate their knowledge and learning processes;
(e) ability to collaborate and work in groups;
(f) communication skills;
(g) ability to solve real-world problems;
(h) understanding of the history of science and the relationship of science to society and other disciplines;
(i) understanding of, or at least appreciation for, "modern" physics;
(j) ability to participate in authentic research.

In my opinion, more support should be given by universities, foundations, and governments to the development of a science of education spearheaded by disciplinary education researchers working in concert with cognitive scientists and education specialists. In the words of cognitive psychologists Anderson et al. (1998),

The time has come to abandon philosophies of education and turn to a Science of Education . . . . . . If progress is to be made to a more scientific approach, traditional philosophies . . . . (such as radical constructivism) . . . . will be found to be like the doctrines of folk medicine. They contain some elements of truth and some elements of misinformation . . . . . . only when a science of education develops that sorts truth from fancy—as it is beginning to develop now—will dramatic improvements in educational practice be seen.

However, it should be emphasized that the development of better strategies for the enhancement of student learning through a Science of Education will not improve the educational system unless (a) university and K–12 teachers are educated to effectively implement those strategies, and (b) universities start to think of education in terms of student learning rather than the delivery of instruction - see e.g., “From Teaching to Learning: A New Paradigm for Undergraduate Education” [Barr & Tagg (1995)], and “The Physics Education Reform Effort: A Possible Model for Higher Education?” Hake (2005a)].

**Acknowledgement**

I thank (a) Josip Slisko of the University of Puebla, Mexico, for informing me of the new journal Latin-American Journal of Physics Education (LAJPE) <http://journal.lapen.org.mx/>, and suggesting that I contribute this article; and (b) LAJPE editor César Eduardo Mora Ley for his helpfulness.
References & Footnotes [Tiny URL’s courtesy <http://tinyurl.com/create.php>. All URL’s were accessed on 18 September 2007.]


Hake, R. R. 2005a. “The Physics Education Reform Effort: A Possible Model for Higher Education?” online at <http://www.physics.indiana.edu/~hake/NTLF42.pdf> (100 kB). This is a slightly edited version of an article that was (a) published in the *National Teaching and Learning Forum* 15(1), December, online to subscribers at <http://www.ntlf.com/FTPSite/issues/v15n1/physics.htm>, and (b) disseminated by the *Tomorrow's Professor* list <http://ctl.stanford.edu/Tomprof/postings.html> as Msg. 698 on 14 Feb 2006. For an executive summary see Hake (2006a).


Klymkowsky, M.W. 2007. "Background on the BCI" online at <http://bioliteracy.net/>, click on "Background on the BCI" (Biology Concept Inventory) in the left column. For other Biology diagnostic tests see Hake (2007c).


Stokstad, E. 2001. “Reintroducing the Intro Course,” Science 293: 1608-1610, 31 August; online at <http://tinyurl.com/266973>. Stokstad wrote: “Physicists are out in front in measuring how well students learn the basics, as science educators incorporate hands-on activities in hopes of making the introductory course a beginning rather than a finale.”


