What Might Psychologists Learn from the Scholarship of Teaching and Learning in Physics? [1,2]

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

Abstract: In this article I:

(a) note that some physicists have been engaged in the "Scholarship of Teaching and Learning" (SoTL) for over four decades;
(b) discuss evidence from SoTL in physics for the approximately two- standard deviation superiority in average pre-to-post-course normalized gains inconceptual understanding for "interactive engagement" over "traditional" passive-student lecture methods;
(c) list some crucial operational definitions;
(d) present accolades from biologists, economists, and mathematicians for SoTL in physics;
(e) discuss an apparent deficiency of SoTL in psychology: the missing "Psychology Concept Inventory";
(f) list 14 hard won lessons from SoTL in physics that may be of value to psychologists;
(g) conclude from all the above that it's conceivable that psychologists might learn something from SoTL in physics.

Key Phrases: concept inventory, interactive engagement, passive-student lecture

---

1. The reference is Hake, R.R. 2014. "What Might Psychologists Learn from the Scholarship of Teaching and Learning in Physics?" submitted to *Scholarship of Teaching and Learning in Psychology* <http://bit.ly/1rzcwUo> on 28 Oct 2014. I welcome comments and suggestions addressed to <rrhake@earthlink.net>.

2. Partially supported by NSF Grant DUE/MDR-9253965.

I. Historical Note

Little known to most of the education community some physicists have been engaged in the "Scholarship of Teaching and Learning" (SoTL) for over four decades. In Section IV, "Empirical Studies," of McDermott & Redish's (1999) "Resource letter on physics education research," I count over 80 examples of SoTL, dating from Joe McKinnon's (1971) "Earth science, density, and the college freshman," wherein he reported the results of testing 143 college freshmen on their understanding of the concept of "density."

One of the earliest examples of SoTL in physics education, appears to be the effort involved in the development of the *Science Curriculum Improvement Study* (SCIS), now available through Delta Education <http://bit.ly/1ou64Mq>.[1] In "One Physicist Experiments with Science Education," Robert Karplus (1964) <<http://bit.ly/1tgf5fc>>[1] wrote:

> "The experimentation with science teaching that I have described is being carried out by the *Science Curriculum Improvement Study* at the University of California in Berkeley. The parts of the science program which have been constructed by SCIS staff members over the past three years are now ready for classroom trial. The kindergarten and first grade teachers in several schools are working with a unit called *Material Objects*, while the second and third grade teachers are working with a unit called *Interaction and Systems* with their classes. . . . . . . Staff members and consultants are available to evaluate the effectiveness of the teaching program and to help participating teachers in using the materials. Reactions and suggestions from the teachers and the results of observations of the pupils' behavior will help determine what revisions in the teaching plans are necessary."

II. Evidence for the Superiority of "Interactive Engagement" Over "Traditional" Passive-student Lecture Methods

More recently, for Newtonian mechanics, physics education researchers have demonstrated that "interactive engagement" (IE) methods can produce a roughly two-standard-deviation superiority in average normalized pre-to-post-course learning gains <g> over "traditional" (T) passive-student lecture methods [Hake (1998a,b)]. As of 2008, similar differences in <g> between IE and T courses had been reported in at least 25 other peer reviewed publications, as listed in "Design-Based Research in Physics Education Research: A Review" [Hake (2008)].

That research involves the measurement of pre-to-post-course learning gains on valid and consistently reliable multiple-choice *Concept Inventories* <<http://bit.ly/dARkDY>> developed by disciplinary experts - see, e.g.: Halloun & Hestenes (1985a,b); Hestenes et al. (1992); Thornton & Sokoloff (1998)] - and the use of reasonably well-matched control groups provided by traditional introductory courses. For reviews of such research see e.g.,

A. "Lessons from the physics education reform effort" [Hake (2002a)]. The abstract reads:

> "A survey of pre/post test data using the Halloun-Hestenes Mechanics Diagnostic testor more recent Force Concept Inventory is reported for 62 introductory physics courses enrolling a total number of students N = 6542. A consistent analysis over diverse student populations in high schools, colleges, and universities is obtained if a rough measure of the average effectiveness of a course in promoting conceptual understanding is taken to be the average normalized gain <g>. The latter is defined as the ratio of the actual average gain (%<post> – %<pre>) to the maximum possible average gain (100 –%<pre>). Fourteen "traditional" (T) courses (N = 2084) which

made little or no use of interactive-engagement (IE) methods achieved an average gain <g> T-ave = 0.23 ± 0.04(std dev). In sharp contrast, forty-eight courses (N = 4458) which made substantial use of IE methods achieved an average gain <g> IE-ave = 0.48 ± 0.14 (std dev), almost two standard deviations of <g> IE-ave above that of the traditional courses. Results for 30 (N = 3259) of the above 62 courses on the problem-solving Mechanics Baseline test of Hestenes-Wells imply that IE strategies enhance problem-solving ability. The conceptual and problem-solving test results strongly suggest that the classroom use of IE methods can increase mechanics-course effectiveness well beyond that obtained in traditional practice."

B. "The Physics Education Reform Effort: A Possible Model for Higher Education" [Hake (2005c)]. Therein I wrote (slightly edited) "In sharp contrast to the:

(1) *invalid* (Student Evaluations of Teaching);

(2) *indirect* [*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 "Re: Measuring Teaching Performance" (Hake, 2005d)]; or

(3) *general-ability* measures such as the *Collegiate Leaning Assessment* (CLA) <<http://bit.ly/ZYmqW3>>[1];

is the direct measure of students' higher-level *domain-specific* learning through pre/post testing using (a) valid and consistently reliable tests *devised by disciplinary experts*, and (b) traditional courses as controls. Such pre/post testing, pioneered by economists (Paden & Moyer, 1969) and physicists (Halloun & Hestenes, 1985a,b), is rarely employed in higher education, in part because of the tired old canonical objections recently lodged by Suskie (2004) and countered by Hake (2004a) and Scriven (2004). Despite the nay-sayers, pre/post testing is gradually gaining a foothold in introductory astronomy, economics, biology, chemistry, computer science, economics, engineering, and physics courses [see Hake (2004b,c) for references]. "

C. "Resource Letter ALIP-1: Active-Learning Instruction in Physics" [Meltzer & Thornton (2012)]. Regarding "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses" [Hake (1998a)] the authors wrote:

"Widely cited analysis of test data from thousands of students in dozens of courses indicating the superior effectiveness of active-learning instruction in physics ("interactive engagement") in comparison to traditional, lecture-based methods."

D. *Adapting to a Changing World - Challenges and Opportunities in Undergraduate Physics Education* [NRC (2013)]. The report states:

"Hake's (1998a) seminal report on the effectiveness of interactive engagement methods remains an important contribution to undergraduate physics education. The article presents results from the Mechanics Diagnostic (MD) (Halloun and Hestenes, 1985) and its successor, the Force Concept Inventory (FCI) (Hestenes et al., 1992), given before and after instruction on Newtonian mechanics in a variety of courses taught using different approaches. . . . . . . the conclusion, that more effective instructional approaches involve active learning, has been supported by many other studies using different methodology [Meltzer and Thornton (2012), Hoellwarth et al. (2005)].

E. "Teaching and physics education research: bridging the gap" [Fraser et al. (2014)]. The authors wrote:

"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. . . . . 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 (1998a)."

But Wait! Can multiple-choice tests (MCTs) measure conceptual understanding and higher-order learning? [For a cogent discussion of "higher-order learning" see Shavelson & Huang (2003).] 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 et al. (1992)] . . . is an assessment that uses multiple-choice items to tap into higher-level cognitive processes."

III. Crucial Operational Definitions [2]

Some definitions [Hake (1998a,b)] are in order:

A. "Interactive engagement" (IE) methods are defined *operationally* as those designed at least in part to promote conceptual understanding through active engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback to both students and instructors through student discussion with peers and/or instructors. The feedback to instructors facilitates *formative evaluation*[3] in the sense used by Black & Wiliam (1998) and Shavelson (2008).

B. "Traditional" (T) methods are defined *operationally* as those which make little or no use of "interactive engagement" methods, relying primarily on passive-student lectures, recipe laboratories (in which detailed and explicit procedures must be followed), and algorithmic problem examinations – this is what's known to most physicists (but not to most cognitive scientists) as "*direct instruction*."

C. The "average normalized gain" <g> is the average *actual* gain [<%post> - <%pre>], divided by the *maximum* possible average actual gain [100% - <%pre>], where the angle brackets <. . . > signify class averages. For a discussion of the rationale and half-century-old history of the "normalized gain" see "Design-Based Research in Physics Education Research: A Review [Hake (2008)].

IV. Accolades from Biologists, Economists, and Mathematicians for SoTL in Physics

A. Biologists Klymkowsky, Garvin-Doxas, & Zeilik (2003) authored an article "Bioliteracy and Teaching Efficiency: What Biologists Can Learn from Physicists." Their abstract reads (slightly edited):

"The introduction of the *Force Concept Inventory* (FCI) [Hestenes et al. (1992)] produced a remarkable impact within the community of physics teachers. An instrument to measure student comprehension of the Newtonian concept of force, the FCI demonstrates that active learning leads to far superior student conceptual learning than didactic lectures.

Compared to a working knowledge of physics, biological literacy and illiteracy have an even more direct, dramatic, and personal impact. They shape public research and reproductive health policies, the acceptance or rejection of technological advances, such as vaccinations, genetically modified foods and gene therapies, and, on the personal front, the reasoned evaluation of product claims and lifestyle choices. While many students take biology courses at both the secondary and the college levels, there is little in the way of reliable and valid assessment of the effectiveness of biological education. This lack has important consequences in terms of general bioliteracy and, in turn, for our society.

Here we describe the beginning of a community effort <http://bioliteracy.colorado.edu/> to define what a bioliterate person needs to know and to develop, validate, and disseminate a tiered series of instruments collectively known as the Biology Concept Inventory (BCI), which accurately measures student comprehension of concepts in introductory, genetic, molecular, cell, and developmental biology. The BCI should serve as a lever for moving our current educational system in a direction that delivers a deeper conceptual understanding of the fundamental ideas upon which biology and biomedical sciences are based."

B. Biologists Wood and Gentile (2003) in a *Science* article "Teaching in a research context" wrote:

"Physics educators have led the way in developing and using objective tests to compare student learning gains in different types of courses, and chemists, biologists, and others are now developing similar instruments. These tests provide convincing evidence that students assimilate new knowledge more effectively in courses including active, inquiry-based, and collaborative learning, assisted by information technology, than in traditional courses."

C. Economists Simkins & Maier (2008) contributed Chapter 36 "Learning from Physics Education Research: Lessons for Economics Education" to the *International Handbook on Teaching and Learning* [Hoyt & McGoldrick (2012)]. That chapter is online at the "Social Science Research Network" (SSRN) at <http://bit.ly/1rr962K> (click on "Download This Paper"). The abstract reads (slightly edited):

"We believe that economists have much to learn from educational research practices and related pedagogical innovations in other disciplines, in particular physics education. In this paper we identify three key features of physics education research that distinguish it from economics education research:

(1) the intentional grounding of physics education research in learning science principles;

(2) a shared conceptual research framework focused on how students learn physics concepts;

(3) a cumulative process of knowledge-building in the discipline;

and describe the influence of the above on new teaching pedagogies, instructional activities, and curricular design in physics education.

In addition, we highlight four specific examples of successful pedagogical innovations drawn from physics education - context-rich problems, concept tests, just-in-time teaching, and interactive lecture demonstrations - and illustrate how these practices can be adapted for economics education."

D. **Mathematician** David Bressoud, in his Mathematics Association of America (MAA) "Launchings" column of 1 July 2012 titled "Learning from the Physicists," wrote:

"The Physics Education Research (PER) community, through the American Association of Physics Teachers, has done a nice job of organizing a website <http://perusersguide.org/> of 51 'Evidence-based teaching methods' that have been demonstrated to be effective. The site is organized to make it useful for the instructor: a brief description and each method and six searchable cross-listings that describe:

(1) <u>Level</u>: the courses for which it is appropriate, usually introductory physics;

(2) <u>Setting</u>: whether designed for large lecture, small classes, labs, or recitation sections;

(3) <u>Coverage</u>: whether it requires studying fewer topics at greater depth;

(4) <u>Effort</u>: low, medium, or high;

(5) <u>Resources</u>: what is needed, from computer access to printed materials that must be purchased, to classrooms with tables;

(6) <u>Skills</u>: what students are expected to acquire, usually including conceptual understanding, but also possibly problem-solving skills and laboratory skills.

In addition, each of the methods includes a list of the types of validation that have been conducted: what aspects of student learning were studied, what skills the method has been demonstrated to improve, and the nature of the research methods . . . .

Unfortunately, the experience of the physicists demonstrates that the existence of research based instructional strategies together with documentation of their effectiveness is not sufficient to guarantee their widespread adoption. Why not? . . . .

[Many of the reasons] can be found in the article "The Use of Research-Based Instructional Strategies in Introductory Physics: Where do Faculty Leave the Innovation-Decision Process [Henderson, Dancy, & Niewiadomska-Bugaj (2002)]. . . . .

The work that they have done via surveys of physics faculty demonstrates that the greatest problem is not in making faculty aware of what has been done, or even in getting faculty to try different approaches to teaching. *The greatest problem is in getting faculty to stick with these strategies*. [My *italics*.]

V. An Apparent Deficiency of SoTL in Psychology: The Missing "Psychology Concept

Inventory”

It’s common knowledge that psychologists have made important contributions to the psychology of learning. For example (in alphabetical order by last names, followed by Wikipedia entries[4])

Richard Atkinson <<http://bit.ly/1F4GQ0h>>[1],
David Ausubel <<http://bit.ly/1w9iixb>>,
David Berliner <<http://bit.ly/1nnkq4H>>,
Hilda Borko <<http://bit.ly/1uaJIjC>>,
John Bransford <<http://bit.ly/1yFo7qD>>,
Ann Brown <<http://bit.ly/1sKUGQm>>,
Jerome Bruner <<http://bit.ly/1rd8b6d>>,
Allan Collins <<http://bit.ly/1vA59jV>>,
Howard Gardner <<http://bit.ly/1w3ne7G>>,
Diane Halpern <<http://bit.ly/JMdE6s>>,
Larry Hedges <<http://bit.ly/1xVFJvi>>,
William James <<http://bit.ly/1qwqyne>>,
Fred Keller <<http://bit.ly/1sVISLf>>,
David Klahr <<http://bit.ly/OdsUKa>>,
David Kolb <<http://bit.ly/1nmhcyp>>,
Marcia Linn <<http://bit.ly/1ov9h2K>>,
Roy Pea <<http://bit.ly/11uRLAZ>>,
Jean Piaget <<http://bit.ly/1yEp0j7>,
Marlene Scardamalia <<http://bit.ly/1rAq9zw>,
Richard Shavelson <<http://bit.ly/VvmqKa>>,
Robert Slavin <<http://bit.ly/1vitxFk>>,
Robert Sternberg <<http://bit.ly/1a4GqFB>>,
Deborah Stipek <<http://bit.ly/11uT5Ux>>, and
Lev Vygotsky <<http://bit.ly/11oBlKG>>.

In addition, in the early 1980s psychologist Michael McCloskey’s <http://bit.ly/1CDobVs> seminal studies of misconceptions regarding “bodies in motion” startled physics educators by showing that even after an introductory course on Newtonian mechanics many students retained naive beliefs about motion. In their pioneering paper “The initial knowledge state of college physics” in which they introduced the landmark “Mechanics Diagnostic Test” of conceptual understanding of Newtonian mechanics, physicists Halloun & Hestenes (1985a) paid tribute to the foundational work of McCloskey and his collaborators, citing “Curvilinear Motion in the Absence of External Forces: Naïve Beliefs About the Motion of Objects” [McCloskey, Caramazza, & Green (1980)] and “Naive beliefs in ‘sophisticated’ subjects: Misconceptions about trajectories of objects” [Caramazza, McCloskey, & Green (1981)].

McCloskey’s work also had an impact on his fellow psychologists. Profitt & Caudek in Chapter 8, “Depth Perception and Perception of Events,” pp. 213-236 of *Handbook of Psychology, Experimental Psychology, Volume* 4 [Healy et al. (2003)], reference “Intuitive Physics” [McCloskey (1983)], “Curvilinear Motion in the Absence of External Forces: Naïve Beliefs About the Motion of Objects” [McCloskey, Caramazza, & Green (1980)], and “Intuitive physics: The straight down belief and its origin” [McCloskey, Washburn, & Felch

(1983)].

Despite the above outstanding contributions by psychologists, contemporary psychology education researchers (PsERs) have evidently *not* developed a “Psychology Concept Inventory” (PCI) comparable to the prototype “Force Concept Inventory” (FCI) [Hestenes et al. (1992)] developed by the physics education researchers (PhERs). My cursory literature search located only two tests of psychology understanding that might be given to introductory course students: E.D. Vaughan's (1977) “Test of Common Beliefs” and L.E. McCutcheon’s (1991) “New Test Of Misconceptions About Psychology.” But these are evidently rarely employed by psychologists to evaluate their teaching. At the very least, since psychology it often taken by incoming university students to fulfill a *science* requirement, it might be interesting for psychologists to administer a “Nature of Science Assessment” [Lederman et al. (2014)] as a pre- and post-test.

Despite the dearth of formative[5] evaluation of introductory psychology course students, there has been some discussion of such evaluation in, e.g., (a) “Confronting Psychological Misconceptions in the Classroom: Challenges and Rewards” [Lilienfeld (2010)]; (b) “College Students’ Common Beliefs about Psychology” [Lamal (1979)]; (c) the 14-post PhysLrnR discussion-list thread “Re: Pretest and Post-test instruments for an Introduction to Psychology" in the January 2007 archives of PhysLrnR at <http://bit.ly/1vMth31> – see e.g., Hake (2007b,c), Holton (2007), and Wittmann (2007)]; and (d) “Why Don't Psychologists Research the Effectiveness of Introductory Courses?” [Hake (2007d)]. In that post I wrote (slightly edited):

> “I argue that there's nothing to prevent the development in *any* discipline of “Interactive Engagement” methods similar to those found in physics to yield average normalized gains about two standard deviations greater than those produced by traditional passive-student lecture courses. It has been suggested by Holton (2007) that the apparent failure of psychologists to engage in such development, even despite their long history of major contributions to education . . . . . . [[see above]]. . . . . can be attributed to a lack of funding. But I would guess (please correct me if I’m wrong) that over the past 20 years there’s been more funding for educational research by psychologists than by physicists. In my opinion, the lack of substantive discussion in the psychology literature of the measurement and enhancement of learning gains in introductory psychology courses suggests that psychologists, with a few notable exceptions, are simply not interested in gauging the effectiveness of their undergraduate courses in enhancing student higher-order learning.”

In “Should We Measure Change? Yes!” [Hake (2011b), I wrote (slightly edited):

> “Formative[5] pre/post testing is being successfully employed to improve the effectiveness of courses in undergraduate astronomy, biology, chemistry, economics, engineering, geoscience, and physics. *But such testing is still anathema to many members of the Psychology-Education-Psychometric (PEP) community*. I argue that this irrational bias impedes a much needed enhancement of student learning in higher education.” . . . .

And in a recent *Carnegie Perspective*, psychologist Lloyd Bond (2005) <<http://bit.ly/1wE0qLa>>, a senior scholar at the Carnegie Foundation, commented (slightly edited):

"If one wished to know what knowledge or skill Johnny has acquired over the course of a semester, it would seem a straightforward matter to assess what Johnny knew at the beginning of the semester and reassess him with the same or equivalent instrument at the end of the semester. It may come as a surprise to many that measurement specialists have long advised against this eminently sensible idea. Psychometricians don't like 'change' or 'difference' scores in statistical analyses because, among other things, they tend to have lower reliability than the original measures themselves. Their objection to change scores is embodied in the very title of a famous paper 'How we should measure 'change'- or should we?' [Cronbach & Furby (1970)]. They opined:

> 'gain scores are rarely useful, no matter how they may be adjusted or refined. . . .investigators who ask questions regarding gain scores would ordinarily be better advised to frame their questions in other ways.' "

Cronbach & Furby's dour appraisal of pre/post testing has echoed down though the literature to present day texts on assessment such as that by Suskie (2009). . . .[[for an antidote to anti-pre/post Suskie (2009) see pro-pre/post Maki (2010)]]. . . . . 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 [Hake (2005c)], is one reason for the glacial progress of educational research [Lagemann (2000, 2002); Berliner (2002)] and the widely acknowledged need for reform of higher education– see e.g.:

(1) *Academically Adrift: Limited Learning on College Campuses* [Arum & Roksa (2011)];
(2) *The American University: How It Runs, Where It Is Going* [Barzun (1968, 1993)];
(3) *Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They Should Be Learning More* [Bok (2005a)];
(4) "Are colleges failing? Higher ed needs new lesson plans" [Bok (2005b)];
(5) *College: What It Was, Is, and Should Be* [Delbanco (2012)],
(6) *A University for the 21st Century* [Duderstadt (2000)];
(7) *Higher Education?: How Colleges Are Wasting Our Money and Failing Our Kids – and What We Can Do About It* [Hacker & Dreifus (2010)];
(8) *Declining by Degrees: Higher Education at Risk* [Hersh & Merrow (2005)];
(9) "What Does College Teach? It's time to put an end to 'faith-based' acceptance of higher education's quality" [Hersh (2005];
(10) *We're Losing Our Minds: Rethinking American Higher Education*" [Keeling & Hersh (2011)];
(11) *Academic Duty* [Kennedy (1999);
(12) *College Unbound: The Future of Higher Education and What It Means for Students* [Selingo (2013)].

Despite the "Accolades from Biologists, Economists, and Mathematicians for SoTL in Physics" (Sect. IV above), psychologists have, for the most part, ignored SoTL in physics. An exception is Chapter 15 "Signature Pedagogies in Introductory Physics" by physicist Mark J. Lattery <http://bit.ly/1rd3CsK> in the exemplary *Exploring Signature Pedagogies: Approaches to Teaching Disciplinary Habits of Mind*, by psychologists Gurung, Chick, & Haynie (2008).

As for the unreliability of 'change scores,' such charges by Lord (1956, 1958) and Cronbach & Furby (1970) have been called into question by . . . . . .[[see "Should We Measure Change?

Yes!" [Hake (2011b)] for the references in the next two sentences]]. . . . . e.g., Rogosa, Brandt, & Zimowski (1982), Zimmerman & Williams (1982), Rogosa & Willett (1983, 1985), Rogosa (1995), Wittmann (1997), Zimmerman (1997), & Zumbo (1999). Furthermore, the measurement of change is an active area of current research by psychometricians such as Collins and Horn (1991), Collins & Sayer (2001), Singer & Willett (2003), Lissitz (2005), and Liu & Boone (2006). All this more recent work should serve as a caution for those who dismiss measurements of change.

In my discussion list post "Do Psychologists Research the Effectiveness of Their Own Introductory Courses? Hake Responds to Sternberg" [Hake (2005b)], I quoted psychologist Jerry Rudmann as follows:

> "The American Psychological Association has been supporting the peer reviewed, scholarly journal, *Teaching of Psychology* <http://top.sagepub.com/> for over 30 years. This publication, of which each new issue appears quarterly, is packed with empirical evaluations done of specific teaching techniques and strategies, psychology courses, and psychology programs."

I responded to Rudmann thusly:

> "Physicist David Meltzer kindly sent me 'Annotated Bibliographies for the *Teaching of Psychology'* for 1999, 2000, and 2001 [Johnson & Schroeder (2000, 2001, 2002)]. A quick scan indicated nothing at all comparable to the rigorous pre/post testing assessments of introductory courses taking place in physics using: (a) research-based tests widely recognized as both valid and reliable, (b) reasonably well-matched control groups (the traditional courses), and (c) experiments in hundreds of different courses with various types of students, teachers, and institutions. Such pre/post testing does not meet the U.S. Dept. of Education's (USDE's) *pseudo* 'gold standard' of randomized control trials – see e.g., '17 Statements by Gold Standard Skeptics' [Hake (2010a))] - but would nevertheless probably pass muster at the USDE's 'What Works Clearing House' <http://ies.ed.gov/ncee/wwc/> as quasi-experimental studies [Shadish et al. (2002)] of especially strong design.
>
> IMHO, without such mass pre/post testing there can be little understanding of the need for, or the results of, various types of reform pedagogies or curricula. I regard the apparent failure of psychologists to research the effectiveness of their own introductory courses as an important issue in education research because, among other things:
>
> (1) One might expect psychologists with their long history of education research [Berliner (1992), Lagemann (2000), Freedheim &Weiner (2004)] and their leading role in classroom-oriented "Design Based Research" <http://bit.ly/mfy8gB> [Kelly (2003); Hake (2008); Kelly, Lesh, & Baek (2008)] to be in the vanguard of those actively researching the effectiveness of own courses and thus serving as role models for other faculty.
>
> (2) Educational psychologists often staff the "Teaching and Learning Centers" of U.S. universities and thus might (but generally do not) influence faculty to research the cognitive effectiveness of their courses through valid and consistently reliable diagnostic tests developed by disciplinary experts, rather than through the usual problematic [Hake (2002b)] student evaluations.

(3) Psychologists and psychometricians seem to be in control of the U.S. Dept. of Education's "What Works Clearinghouse" <http://ies.ed.gov/ncee/wwc/> and NCLB <http://www2.ed.gov/nclb/landing.jhtml>. Why should they be the arbiters of "What Works" when, as far as I know, they haven't even bothered to research "What Works" in their own courses?"

VI. Lessons From SoTL in Physics

As indicated in "Lessons from the Physics Education Reform Effort" [Hake (2002a)]:

"For more than three decades, physics education researchers have repeatedly shown that traditional introductory physics courses with passive student lectures, recipe labs, and algorithmic problem exams are of limited value in enhancing students' conceptual understanding of the subject [McDermott and Redish (1999)]. Unfortunately, this work was largely ignored by the physics and education communities 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, Halloun et al., 1995) 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 (and thus can be given as an introductory course pretest), yet at the same time are rigorous enough for the initiate."

Here are 14 hard won lessons derived from "Lessons from the Physics Education Reform Effort" [Hake (2002a)] and "Six Lessons From the Physics Education Reform Effort" [Hake (2007c)]. I suspect that these lessons might be beneficial to some in the psychology community.

A. Six Lessons on Interactive Engagement

(1) The use of "Interactive Engagement" (IE) strategies can increase the effectiveness of conceptually difficult courses well beyond that obtained by traditional passive-student lecture methods.

(2) The use of IE and/or high-tech methods, by themselves, does not ensure superior student learning.

(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.

(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.

(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.

(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.

B. Eight Lessons on Implementation

(7) Teachers who possess both content knowledge and "pedagogical content knowledge" are more apt to deliver effective instruction.

(8) College and university faculty tend to overestimate the effectiveness of their own instructional efforts and thus tend to see little need for educational reform.

(9) Such complacency can sometimes be countered by administering high-quality standardized tests of understanding and by "video snooping."

(10) A major problem for undergraduate education in the United States is the inadequate preparation of incoming students, in part due to the inadequate university education of K–12 teachers.

(11) Interdisciplinary cooperation of instructors, departments, institutions, and professional organizations is required for synthesis, integration, and change in the entire chaotic educational system (b) for a compilation of references on "systems thinking" see "Over Two Hundred Annotated References on Systems Thinking" [Hake (2009b).

(12) Various institutional and political factors, including the culture of research universities slow educational reform . . . .[[see "Changing the Culture of Science Education at Research Universities (Anderson et al. (2011) and "Should the Culture of University Science Education Be Changed" (Hake, 2011a)]]. . . . . .

(13) The monumental inertia of the educational system may thwart long-term national reform.

(14) "Education is not rocket science, it's much harder."

- George (Pinky) Nelson, astronaut, astrophysicist, and former director of the AAAS Project 2061, as quoted by E.F. (Joe) Redish (1999).

VII. Conclusions

Based on all the above, I think it's conceivable that psychologists (like biologists, economists, and mathematicians) might learn something from the Scholarship of Teaching and Learning in physics.

VIII. Acknowledgments

This work received partial support from National Science Foundation Grant DUE/MDR9253965.

I have benefited from conversations or correspondence with Patricia Cohen, Robert DeHaan, Jerry Epstein, Robert Fuller, Paul Hickman, Brett Magill, David Meltzer, Robert Mislevy, Jeanne Narum, Ed Nuhfer, Antti Savinainen, Lee Sechrest, Werner Wittmann, LarryWoolf, Donald Zimmerman, and Bruno Zumbo. In addition, my ideas on education have benefited from exchanges on discussion lists in diverse disciplines [see e.g., the listing in “A Guide to the ADLsphere: Over Eighty Academic Discussion Lists With Links to Their Archives & Search Engines” [Hake (2010b)]. Most importantly I thank psychologist Regan A.R. Gurung for suggesting that I write this article and for his valuable comments.

Notes

1. Most URLs are shortened by <http://bit.ly/> and surrounded by angle brackets <....> so as to indicate what is and is not part of the URL and to promote linking of URLs which cross line breaks. Double angle brackets <<....>> surround URLs for Wikipedia articles.

2. “When we say force is the cause of motion we talk metaphysics, and this definition, if we were content with it, would be absolutely sterile. For a definition to be of any use, it must teach us to measure force; moreover, that suffices; it is not at all necessary that it teach us what force is in itself, nor whether it is the cause or the effect of motion.” - Henri Poincaré (1905)

3. Here *formative evaluation* means “all those activities undertaken to provide information to be used as feedback so as to adapt the teaching to meet student needs” as distinguished from “formative evaluation” as evaluation “designed and used to improve an intervention, especially when it is still being developed” - see “Two Different Meanings of 'Formative Evaluation' #2” [Hake (2014)].

4. Those who dismiss Wikipedia entries as a mere “opinion pieces,” may not be aware that a study by *Nature* [Giles (2005)] indicated that *Wikipedia* comes close to *Britannica* in terms of the accuracy of its science entries – see e.g. “In Defense of Wikipedia” [Hake (2009)] at <http://bit.ly/fb4bJx>.

5. Here *formative evaluation* means evaluation “designed and used to improve an intervention, especially when it is still being developed” as distinguished from “formative evaluation” in the sense of “all those activities undertaken to provide information to be used as feedback so as to adapt the teaching to meet student needs”– see “Two Different Meanings of 'Formative Evaluation' #2” [Hake (2014)].

References and Footnotes [All URLs accessed on 28 Oct 2014; most shortened by <http://bit.ly/>.

Note: A few references are to posts on the CLOSED! archives <http://bit.ly/nG318r> of the physics education discussion list PhysLrnR. To access the archives of PhysLnR one needs to subscribe **: - (** , but that takes only a few minutes by clicking on <http://bit.ly/nG318r> and then clicking on “Join or Leave PHYSLRNR-LIST.” 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!

Alexander, P.A. & P.H. Winne, eds. 2006, 2009. *Handbook of Educational Psychology*, 2nd ed. Routledge, publisher’s information at <http://bit.ly/1ntfFqx>. Amazon.com information at <http://amzn.to/1Dkl0UZ>. An expurgated Google book preview of the 2006 printing is online at

<http://bit.ly/1ntflI7>.

Anderson, W.A., U. Banerjee, C.L. Drennan, S.C.R. Elgin, I.R. Epstein, J. Handelsman, G.F. Hatfull, R. Losick, D.K. O'Dowd, B.M. Olivera, S.A. Strobel, G.C. Walker, I.M. Warner. 2011. "Changing the Culture of Science Education at Research Universities," *Science*, 14 January: **331**(6014): 152-153; online as a 172 kB pdf at <http://bit.ly/eSLoCl>; supporting online references suggested by R.R. Hake are at <http://bit.ly/g24Iqm>.

Arum, R. & J. Roksa. 2011. *Academically Adrift: Limited Learning on College Campuses*. University of Chicago Press, publisher's information, including a synopsis and bio, are online at <http://bit.ly/RrkWhK>. Amazon.com information at <http://amzn.to/f1f45O>.

Banta, T.W. & C. Blaich. 2011. "Closing the Assessment Loop," *Change Magazine*, January/February; online at <http://bit.ly/lQyEYp>.

Barzun, J. 1968, 1993. *The American University: How It Runs, Where It Is Going*, University Of Chicago Press; 2nd edition, publisher's information at <http://bit.ly/10umFkS>. First published in 1968. Amazon.com information at <http://amzn.to/126DpRC>, note the searchable "Look Inside" feature. An expurgated Google book preview is online at <http://bit.ly/13tFUjR>.

Benezet, L.P. 1935/36. "The teaching of arithmetic I, II, III: The story of an experiment," *Journal of the National Education Association* **24**(8): 241-244 (1935); **24**(9): 301-303 (1935); **25**(1): 7-8 (1936). The articles were: (a) reprinted in the *Humanistic Mathematics Newsletter* **6**: 2-14 (May 1991); (b) placed on the web along with other Benezetia at the Benezet Centre <http://bit.ly/926tiM>. See also Mahajan & Hake (2000).

Berliner, D.C. 1992. "The science of psychology and the practice of schooling: The one hundred year journey of educational psychology from interest, to disdain, to respect for practice," in Fagan & VandenBos (1992). I've not found an online version of this valuable article, but some of its content is available in Berliner's (2006, 2009) chapter "Educational Psychology: Searching For Essence Throughout a Century of Influence" in the Google book preview <http://bit.ly/1ntflI7> of Alexander & Winne's (2006, 2009) *Handbook of Educational Psychology*, 2nd edition. Berliner (1992) wrote [slightly edited, my *italics*; my insert at ". . . . .[[insert]]. . . ."]:

> "One of those who set the stage for Thorndike was the great muckraker and classroom observer Joseph Mayer Rice (1857- 1934) <http://bit.ly/1sSgYjc>, *the father of research on teaching*. Rice endured great difficulties for his beliefs just a few years before the experimental psychology of E. L. Thorndike <http://bit.ly/1ofcrHD> was deemed acceptable . . . . . In 1897, in Atlantic City, New Jersey, Rice was asked to present his empirical classroom-based research on the futility of the *spelling grind* to the annual meeting of school superintendents. I do not think they were as polite as today's administrators, as they attacked the speaker, yelling the equivalent of "give him the hook." .
> . . . . .[[For the futility of the early-grade of *algorithmic math grind* see Benezet (1935/36)] - for a "give him the hook" response see e.g., Wayne Bishop's (2002) Math-Teach post "Benezet experiment described in the Manchester paper,". . . . ." at <http://bit.ly/1sSfZj0>.]].
> . . . . . Leonard P. Ayres <<http://bit.ly/1FeH19m>> reports on the meeting as follows:

‘The presentation of these data threw that assemblage into consternation, dismay, and indignant protest. *But the resulting storm of vigorously voiced opposition was directed, not against the methods and results of the investigation, but against the investigator who had pretended to measure the results of teaching spelling by testing the ability of the children to spell.’* ”

Berliner, D.C. 2002. “Educational research: The hardest science of all,” *Educational Researcher* **31**(8): 18-20; online as a 70 kB pdf at <http://bit.ly/ZGJRDE>.

Berliner, D. 2006. “Educational Psychology: Searching For Essence Throughout a Century of Influence,” Chapter 1, pp. 3-28 in *Handbook of Educational Research*, 2nd ed. [Alexander & Winne (2006, 2009)]. Most of Berliner’ chapter can be accessed at Amazon.com’s “Look Inside” feature for Alexander & Winne (2006, 2009)].

Bishop, W. 2002. “Benezet experiment described in the Manchester paper,” online on the OPEN! Math-Learn archives at <http://bit.ly/1sVqimw>. Post of 9 Feb 9, 2:40am to Math-Learn. Bishop wrote:

> “The ‘scholarly articles’ that L. P. Benezet published in 1935 in the *Journal of the National Education Association* <http://bit.ly/QP4uGW> are interesting in that they show how little the adjective ‘scholarly’ meant in professional education then, just as how little it means now. Little has changed in the intervening decades. A pied-piper ‘researcher’ comes with a conviction and the gift of leadership, ignores parental concerns especially those of educated parents, finds a way to implement the experiment in spite of the concerns of knowledgeable parents, and goes to his grave knowing that he’s right in spite of the utter demise of his experiment. Amazing. ”

Black, P. & D. Wiliam. 1998. “Inside the Black Box: Raising Standards Through Classroom Assessment,” *Phi Delta Kappan* **80**(2): 139-144, 146-148; online as a 45 kB pdf at <http://bit.ly/1qDZaDZ>.

Bok, D. 2005a. *Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They Should Be Learning More*. Princeton University Press, publisher’s information at <http://bit.ly/qfNIWM>, including the “Table of Contents” and the “Introduction.” Amazon.com information at <http://amzn.to/1sNosUC>, note the searchable “Look Inside” feature.

Bok, D. 2005b. “Are colleges failing? Higher ed needs new lesson plans,” *Boston Globe*, 18 December, online at <http://bo.st/ffRQFa>,

Bond. L. 2005. “A different way to think about teaching and learning - Who has the lowest prices,” *Carnegie Perspectives,* this will presumably eventually appear on the *Carnegie Perspectives* archives at <http://bit.ly/ZZvxWH>.

Bressoud, D. 2012a. “Learning from the Physicists,” MAA “Launchings” column, 1 July 2012; online at <http://bit.ly/MrAuyZ >.

Caramazza, A., M. McCloskey, & B. Green. 1981. “Naive beliefs in ‘sophisticated’ subjects:

Misconceptions about trajectories of objects," *Cognition* **9**: 117; an abstract, online at <http://bit.ly/1rlGnNd>, reads:

> "University students were asked to solve simple problems about the trajectories of falling objects. A majority of the students revealed a variety of misconceptions about motion. However, the few basic patterns of responses produced by the subjects suggest considerable commonality in the types of naive physical "laws" people develop on the basis of everyday experience with the world."

Chick, N.L., A. Haynie, & R.A.R. Gurung, eds., 2012. *Exploring More Signature Pedagogies: Approaches to Teaching Disciplinary Habits of Mind*. Stylus, foreword by Anthony A. Ciccone. publisher's information at <http://bit.ly/1sYa5Oz>. Amazon.com information at <http://amzn.to/1ql7sQU>, note the searchable "Look Inside" feature.

Cronbach, L. & L. Furby. 1970. "How we should measure 'change'- or should we?" *Psychological Bulletin* **74**: 68-80. An abstract, online at <http://bit.ly/LqCSHy> reads:

> "Examines procedures previously recommended by various authors for the estimation of 'change' scores, 'residual,' or 'basefree' measures of change, and other kinds of difference scores. A procedure proposed by F. M. Lord is extended to obtain more precise estimates, and an alternative to the L.R. Tucker, F. Damarin, and S.A. Messick procedure is offered. A consideration of the purposes for which change measures have been sought in the past leads to a series of recommended procedures which solve research and personnel-decision problems without estimation of change scores for individuals."

Delbanco, A. 2012. *College: What It Was, Is, and Should Be*. Princeton University Press, publisher's information at <http://bit.ly/1n89tTC>. Amazon.com information at <http://amzn.to/KmIvkY>.

Duderstadt, J.J. 2000. *A University for the 21st Century*. Univ. of Michigan Press, publisher's information at <http://bit.ly/1od73Fa>. Amazon.com information at <http://amzn.to/lxT8YU>, note the searchable "Look Inside" feature. See also Duderstadt (2001).

Duderstadt, J. J. 2001. "Science policy for the next 50 years: from guns to pills to brains," in *Proceedings of the AAAS Annual Meeting, San Francisco*, February; online at <http://bit.ly/11wV9vo>.

Fagan, T.K. & G.R. VandenBog, eds. 1992. *Exploring applied psychology: Origins and critical analysis: Master lecturers*, American Psychological Association, publisher's information <http://bit.ly/1typ0jB>. Amazon.com information at <http://amzn.to/11xn527>.

Fraser, J.M., A.L Timan, K. Miller, J.E Dowd, L. Tucker, & E. Mazur. 2014. "Teaching and physics education research: bridging the gap." *Rep. Prog. Phys.* **77:** 032401; online at <http://bit.ly/1qITBqi>.

Freedheim, D.K. & I.B. Weiner. 2004. *Handbook of Psychology, History of Psychology, Volume 1, 2nd ed.* Wiley, publisher's information at <http://bit.ly/10aqP9p>. Amazon.com information at <http://amzn.to/1sDcGuD>, note the searchable "Look Inside" feature.

Fuller, R.G., ed. 2002. *A Love of Discovery: Science Education - The Second Career of Robert Karplus*. Springer. Amazon.com information at <http://amzn.to/axnQGn>, note the searchable "Look Inside feature. Author's information is online at <http://bit.ly/1vzaNBv>. This is a valuable resource containing seminal papers of Karplus and his colleagues.

Giles, J. 2005. "Special Report - Internet encyclopaedias go head to head - *Wikipedia* comes close to *Britannica* in terms of the accuracy of its science entries, a *Nature* investigation finds," *Nature* **438**: 900-901, 15 December; online at <http://bit.ly/xxEgXG>.

Goffe, B. 2011. "Re: Business agenda for K-12 STEM education: not research-informed" online on the CLOSED PhysLrnR archives at <http://bit.ly/1we0P8W>. Post of 26 Apr 2011 23:47:19-0400 to PhysLrnR.

Gurung, R.A.R., N.L. Chick, & A. Haynie, eds. 2008. *Exploring Signature Pedagogies: Approaches to Teaching Disciplinary Habits of Mind*, foreword by Anthony A. Ciccone. Stylus. Publisher's information at <http://bit.ly/1EQLeQt>. Amazon.com information at <http://amzn.to/1o8d6e2>, note the searchable "Look Inside" feature. A Google book preview is online at <http://bit.ly/ZBcxhh>. From the foreword by Ciccone, Senior Scholar and Director, Carnegie Academy for the Scholarship of Teaching and Learning:

> "These authors have clearly shown the value in looking for the signature pedagogies of their disciplines. Nothing uncovers hidden assumptions about desired knowledge, skills, and dispositions better than a careful examination of our most cherished practices. The authors inspire specialists in other disciplines to do the same. Furthermore, they invite other colleagues to explore whether relatively new, interdisciplinary fields such as Women's Studies and Global Studies have, or should have, a signature pedagogy consistent with their understanding of what it means to 'apprentice' in these areas."

See also Schwartz & Gurung (2012) and Chick, Haynie, & Gurung (2012).

Hacker, A. & C. Dreifus. 2010. *Higher Education?: How Colleges Are Wasting Our Money and Failing Our Kids - and What We Can Do About It*. Holt/Times Books. Amazon.com information at <http://amzn.to/bunggt>. See also the Hacker/Dreifus blog <http://bit.ly/gJp4Pg>.

Hake, R.R. 1998a. "Interactive-engagement vs traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *Am. J. Phys.* **66**(1): 64-74; online as an 84 kB pdf at <http://bit.ly/9484DG>.

Hake, R.R. 1998b. "Interactive-engagement methods in introductory mechanics courses," online as a 108 kB pdf at <http://bit.ly/aH2JQN>. A crucial companion paper to Hake (1998a).

Hake, R.R. 2002a. "Lessons from the Physics Education Reform Effort" *Ecology and Society* **2**: 28; online at <http://bit.ly/aL87VT>. For an update on the six lessons on "interactive engagement" see Hake (2007a).

Hake, R.R. 2002b. "Re: Problems with Student Evaluations: Is Assessment the Remedy?" online as a 74 kB pdf at <http://bit.ly/hjt9ll>.

Hake, R.R. 2004a. “Re: pre- post testing in assessment,” on the OPEN! ASSESS archives at <http://bit.ly/10o0IvY>. Post of 19 Aug 2004 13:56:07-0700 to ASSESS and various other discussion lists. I respond to the 8 anti-pre/post-testing points made by Linda Suskie (2004) in her ASSESS post of 19 Aug 2004 08:19:53-0400 titled "Re: pre- post testing in assessment."

Hake, R.R. 2004b. “Re: Measuring Content Knowledge,” online on the OPEN! POD archives at <http://bit.ly/c0g8jA>. Post of 14 Mar 2004to POD and various other discussion lists. See also Hake (2004c).

Hake, R.R. 2004c. “Re: Measuring Content Knowledge,” online on the OPEN! POD archives at <http://ntrda.me/1FTwfpr>. Post of 15 Mar 2004 to POD.

Hake, R.R. 2005a. “Do Psychologists Research the Effectiveness of Their Own Introductory Courses?” online on the OPEN! TIPS (Teaching in the Psychological Sciences) archives at <http://bit.ly/1sKhPlY>. TIPS post of 19 Feb 2005 07:58:43-0800. See also Hake (2005b).

Hake, R.R. 2005b. “Do Psychologists Research the Effectiveness of Their Courses? Hake Responds to Sternberg,” online on the OPEN! AERA-L archives at <http://bit.ly/mZaZKc>. Post of 21 Jul 2005 22:55:31-0700 to AERA-L and various other discussion lists.

Hake, R. R. 2005c. “The Physics Education Reform Effort: A Possible Model for Higher Education,” online as a 100 kB pdf at <http://bit.ly/9aicfh>. This is a slightly edited version of an article that was (a) published in the *National Teaching and Learning Forum 15(*1), December 2005, and (b) disseminated by the *Tomorrow's Professor* list <http://bit.ly/d09Y8r> as Msg. 698 on 14 Feb 2006.

Hake, R.R. 2005d. “Re: Measuring Teaching Performance,” online on the OPEN! AERA-L archives at <http://bit.ly/1tB5yAE>. Post of 13 May 2005 17:16:48-0400 to POD and various other discussion lists.

Hake, R.R. 2006. “Possible Palliatives for the Paralyzing Pre/Post Paranoia that Plagues Some PEP’s,” *Journal of MultiDisciplinary Evaluation* **6**, November, online at <http://bit.ly/NuGIRL>. This was a precursor to “Should We Measure Change? Yes!” [Hake (2011b)].

Hake, R.R. 2007a. “Six Lessons From the Physics Education Reform Effort,” *Latin American Journal of Physics Education* **1**(1): 24-31, September; online with AIP style numbered references as initially published at <http://bit.ly//12eL2LS> (scroll down to page 24). Also available with APA style references as a 684 kB pdf at <http://bit.ly/96FWmE>

Hake, R.R. 2007b. “Re: Pretest and Post-test instruments for an Introduction to Psychology,” online on the CLOSED! PhysLrnR archives at <http://bit.ly/1x1PW8p>. Post of 20 Jan 2007 23:01:17-0800 to PhysLrnR. Therein I respond to a psychologist who posted on the PsychTeacher list <http://bit.ly/1rKMvkD> “I would like to see some examples of pretest and post-test measuring devices (knowledge) related to an Introduction to Psychology course.” I responded:

“Asking psychologists for information on pre/post testing is like asking physicists for information on divining rods; see e.g., ‘Do Psychologists Research the Effectiveness of Their Courses? Hake Responds to Sternberg’ [Hake (2005b)]”.

Hake, R.R. 2007c. “Re: Pretest and Post-test instruments for an Introduction to Psychology,” online on the CLOSED! PhysLrnR archives at <http://bit.ly/1y3pzA4>. Post of 23 Jan 2007 22:09:50-0800 to PhysLrnR. Therein I respond to Michael Wittman’s (2007) claim that “. . . . .we can't define Interactive Engagement” – but see the definition of “Interactive Engagement” in the above Sect. III “Crucial Operational Definitions.”

Hake, R.R. 2007d. “Why Don't Psychologists Research the Effectiveness of Introductory Courses?” online on the CLOSED! PhysLrnR archives at <http://bit.ly/10kL9F6>. Post of 24 Jan 2007 15:11:33-0800 to PhysLrnR.

Hake, R.R. 2008. “Design-Based Research in Physics Education Research: A Review,” in Kelly, Lesh, & Baek (2008)]. A pre-publication version of that chapter is online as a 1.1 MB pdf at <http://bit.ly/9kORMZ>.

Hake, R.R. 2009a. “In Defense of Wikipedia,” online on the OPEN ! AERA-L archives at <http://bit.ly/fb4bJx>. Post of 31 Aug 2009 16:41:53-0700 to AERA-L, Net-Gold, and MathTeach.

Hake, R.R. 2009b. “Over Two-Hundred Annotated References on Systems Thinking,” online as a 1.7 MB pdf at <http://bit.ly/9gZdXU >. The abstract and link to the complete post are on my website “Hake's Ed Stuff” at <http://bit.ly/duhqLE> with a provision for comments. For other related posts go to “Hake's Ed Stuff” at <http://HakesEdStuff.blogspot.com/> and scroll down to and click on “Systems Thinking” under “Labels” in the right-hand column.

Hake, R.R. 2010a. "Seventeen Statements by Gold-Standard Skeptics #2," online on the OPEN! AERA-L archives at <http://bit.ly/1xWA7AV>. Post of 27 Apr 2010 10:03:58-0700 to AERA-L and Net-Gold. The abstract and link to the complete post were transmitted to various discussion lists and are online on my blog “Hake’sEdStuff” at <http://bit.ly/bSmKI9> with a provision for comments.

Hake, R.R. 2010b. “Over Eighty Academic Discussion Lists With Links To Their Archives & Search Engines,” online as a 3.7 MB pdf at <http://bit.ly/970OZr>.

Hake, R.R. 2011a. “Should the Culture of University Science Education Be Changed?” Invited talk, Southern California Section of the AAPT, 14 May 2011, Pierce College, Woodland Hills, CA; online as a 3.2 MB pdf at <http://bit.ly/iegznz> and as reference #63 at <http://bit.ly/b2UsK6>. The abstract and link to the complete post are on my blog "Hake'sEdStuff" at <http://bit.ly/lCyN97> with a provision for comments.

Hake, R.R. 2011b. “Should We Measure Change? Yes!” online as a 2.5 MB pdf at <http://bit.ly/d6WVKO> (2.5 MB). To appear as a chapter in "Rethinking Education as a Science" [Hake (in preparation)]. A severely truncated version appears in Hake (2006). The abstract reads:

“Formative pre/post testing is being successfully employed to improve the effectiveness of courses in undergraduate astronomy, biology, chemistry, economics, engineering, geoscience, math, and physics. But such testing is still anathema to many members of the psychology-education-psychometric (PEP) community. I argue that this irrational bias impedes a much needed enhancement of student learning in higher education. I then review the development of diagnostic multiple-choice tests of higher-level learning; normalized gain and ceiling effects; the documented two-sigma superiority of interactive engagement (IE) to traditional passive-student pedagogy in the conceptually difficult subject of Newtonian mechanics; the probable neuronal basis for such superiority; education's lack of a community map; higher education's resistance to change and its related failure to improve the public schools; and, finally, why we should be concerned with student learning."

See also "The Impact of Concept Inventories On Physics Education and Its Relevance For Engineering Education" [Hake (2011c)].

Hake, R.R. 2011c. "The Impact of Concept Inventories On Physics Education and Its Relevance For Engineering Education" invited talk, 8 August, second annual NSF-sponsored "National Meeting on STEM Concept Inventories," Washington, D.C., online as a 8.7 MB pdf at <http://bit.ly/nmPY8F>.

Hake, R.R. 2014. "Two Different Meanings of 'Formative Evaluation' #2," online on the OPEN! AERA-L archives at <http://bit.ly/1e8Zhpr>. Post of 29 Jan 2014 19:06:18-0800 to AERA-L and Net-Gold. The abstract and link to the complete post were transmitted to several discussion lists and are on my blog "Hake'sEdStuff" at <http://bit.ly/MxnMBt> with a provision for comments.

Halloun, I. & Hestenes, D. 1985a. "The initial knowledge state of college physics," *Am. J. Phys.* **53**(11): 1043-1055; online at <http://bit.ly/b1488v>, scroll down to "Evaluation Instruments."

Halloun, I. & D. Hestenes. 1985b. "Common sense concepts about motion." *Am. J. Phys.* **53**(11) 1056-1065; online at <http://bit.ly/b1488v>, scroll down to "Evaluation Instruments."

Halloun, I., R.R. Hake, E.P. Mosca, & D. Hestenes. 1995. "Force Concept Inventory (1995 Revision)," online (password protected) at <http://bit.ly/b1488v>, scroll down to "Evaluation Instruments."

Currently available in 27 languages: Arabic, Chinese (simplified), Chinese (traditional), Croatian, Czech, Dutch, English, Finnish (Suomi), French (Canadian), French (France), German, Greek, Hungarian, Indonesian, Italian (Italiano), Japanese, Malaysian (Bahasa), Norwegian, Persian, Polish Portuguese (Portugal), Portuguese (Brazil), Russian, Slovak, Spanish, Swedish, & Turkish.

Healy, A.F., R.W. Proctor, & I.B. Weiner, eds. 2003. *Handbook of Psychology, Experimental Psychology (Volume 4),* Wiley. Amazon.com information at <http://amzn.to/1vPb5Wv>. An expurgated Google book preview is online at <http://bit.ly/1ojb2zR>.

Hestenes, D., M. Wells, & G. Swackhamer. 1992. "Force Concept Inventory," *Phys. Teach.* **30**(3): 141-158; online as a 100 kB pdf at <http://bit.ly/foWmEb> [but without the test itself]. For the 1995 revision see Halloun et al. (1995).

Henderson, C., M. Dancy, and M. Niewiadomska-Bugaj. 2012. “The Use of Research-Based Instructional Strategies in Introductory Physics: Where do Faculty Leave the Innovation-Decision Process,” *Physical Review Special Topics - Physics Education Research* **8** (2), online as a 938 kB pdf at <http://bit.ly/1nsuyt2>.

Hersh, R.H. 2005. “What Does College Teach? It’s time to put an end to ‘faith-based’ acceptance of higher education’s quality,” *Atlantic Monthly* **296**(4): 140-143, November; freely online to (a) subscribers of the *Atlantic Monthly* at <http://theatln.tc/1iRsDMD>, and (b) (with hot-linked academic references) to educators at <http://bit.ly/ZEP4LI> (scroll to the APPENDIX). See also Hersh & Merrow (2005).

Hersh R.H. & J. Merrow, eds. 2005. *Declining by Degrees: Higher Education at Risk*. Palgrave Macmillan, publisher’s information at <http://bit.ly/1s3prLq> Amazon.com information at <http://amzn.to/ftlof5>.

Hoellwarth, C., M.J. Moelter, & R.D. Knight, 2005. “A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms,” *Am. J. Phys*. **73**(5): 459-463; online to subscribers at <http://bit.ly/156hHhF>.

Holton, D. 2007. “Re: Pretest and Post-test instruments for an Introduction to Psychology,” PhysLrnR post of 20 Jan 2007 23:01:17-0800; online at <http://bit.ly/1rQCYWl>.

Hoyt, G.M. & K.M. McGoldrick. 2012, *International Handbook on Teaching and Learning Economics*. Edward Elgar, publisher’s information at <http://bit.ly/1vo9err>. Amazon.com information at <http://amzn.to/1tXd1WP>, note the searchable “Look Inside” feature. An expurgated Google book preview is online at <http://bit.ly/1njGRI9>.

Johnson, D.E., & S.I. Schroder. 2000. “Annotated Bibliography on the Teaching of Psychology: 1999,” *Teaching of Psychology* **27**(4); 296-303; a first page preview is online at <http://bit.ly/1uzWR67>.

Johnson, D.E., & S.I. Schroder. 2001. “Annotated Bibliography on the Teaching of Psychology: 2000,” *Teaching of Psychology* **28**(4); 303-309; a first page preview is online at <http://bit.ly/1zjm9NY>.

Johnson, D.E., & S.I. Schroder. 2002. “Annotated Bibliography on the Teaching of Psychology: 2001,” *Teaching of Psychology* **29**(4); 337-344; a first page preview is online at <http://bit.ly/ZEMTrv>.

Karplus. R. 1964. “One Physicist Experiments with Science Education,” *Am. J. Phys*. **32**(11): 837-839; online to subscribers at <http://bit.ly/OYk9FD>. For reprints of Karplus’s pioneering educational research and development work, dating from 1962, see Fuller (2002).

Keeling, R.P. & R.H. Hersh. 2011. *We're Losing Our Minds: Rethinking American Higher Education,*" Palgrave Macmillan, publisher's information at <http://bit.ly/IOE8wU>. Amazon.com information at <http://amzn.to/HXMxN8>, note the searchable "Look Inside"

Feature. The publisher’s information includes a review at <http://bit.ly/Ied4Z1> by Steven S. Volk, Professor of History and Director of the Center for Teaching Innovation and Excellence, Oberlin College.

Kelly, A.E. 2003. “Research as Design,” *Educational Researcher* **32**(1): 3-4; online as a 901 kB pdf at <http://bit.ly/1t1fiVK>. Introduces an issue devoted to the theme "The Role of Design in Educational Research."

Kelly, A.E., R.A. Lesh, & J.Y. Baek. 2008. *Handbook of Design Research Methods in Education: Innovations in Science, Technology, Engineering, and Mathematics Learning and Teaching*," Routledge. Publisher's information at <http://bit.ly/dkLabI>; Amazon.com information at <http://amzn.to/gtRpbU>, note the searchable “Look Inside” feature. An expurgated Google book preview is online at <http://bit.ly/1sW4urd>.

Kennedy, D. 1999. *Academic Duty*. Harvard University Press, publisher's information at <http://bit.ly/edBogg>. Amazon.com information at <http://amzn.to/hbeAsp>, note the "Look Inside" feature. An expurgated Google book preview is online at <http://bit.ly/hmuC5R>.

Klymkowsky, M.W., K. Garvin-Doxas, & M. Zeilik. 2003. “Bioliteracy and Teaching Efficiency: What Biologists Can Learn from Physicists,” *Cell Biology Education* **2**: 155-161; online at <http://bit.ly/9A1Arx>.

Lagemann, E.C. 2000, 2002. *An Elusive Science: The Troubling History of Education Research*. Univ. of Chicago Press, publisher’s information at <http://bit.ly/1vjfjUE>. Amazon.com information at <http://amzn.to/1phsjEq>.

Lamal, P.A. 1979. “College Students' Common Beliefs about Psychology,” *Teaching of Psychology* **6**(3): 155-158; an APA PsyNet abstract at <http://bit.ly/1CCXDDH> reads:

> “A shortened form of E. D. Vaughan's (1977) ‘Test of Common Beliefs’ was administered to 162 undergraduates at the beginning and end of an introductory psychology course (PC) and to 44 seniors before and after a senior level PC. Results support those of Vaughan, in that an introductory PC seems to have little effect on some common beliefs about the subject that are widely held by students. (10 ref)”

On 17 Oct 2014 a:

(1) Google <http://www.google.com/> search for “College Students’ Common Beliefs about Psychology” yielded 177 hits at <http://bit.ly/1Fe5njA>.

(2) Google Scholar <http://scholar.google.com> for the same phrase indicated 38 citations at <http://bit.ly/1DjAcl3>. These might be worth examining.

Lederman, N.G., S.A. Bartos, & J.S. Lederman. 2014. “The Development, Use, and Interpretation of Nature of Science Assessments,” Chapter 29, pp. 971-997 in *International Handbook of Research in History, Philosophy and Science Teaching* [Matthews (2014)]. The first two pages are online at <http://bit.ly/12jEhbT>. All but 6 pages are online in the Google book preview at <http://bit.ly/1wc5QAq> of Matthews (2014).

Lilienfeld, S.O. 2010. “Confronting Psychological Misconceptions in the Classroom: Challenges and Rewards," Association for Psychological Science, *Observer* **23**(7); online at <http://bit.ly/1yLkAHl>.

Lord, F.M. 1956. "The measure of growth," *Educational and Psychological Measurement* **16**: 42-437. I've not found an online version of this article.
Lord, F.M. 1958. "Further problems in the measurement of growth," *Educational and Psychological Measurement* **18**: 437-454. I've not found an online version of this article.

Mahajan, S. & R.R. Hake. 2000. "Is it time for a physics counterpart of the Benezet/Berman math experiment of the 1930's?" Physics Education Research Conference 2000: Teacher Education; online as a 319 kB pdf at <http://bit.ly/1Fmxa1o> and as article #6, at the Benezet Centre <http://bit.ly/926tiM>.

Maki, P.L. 2010. *Assessing for Learning: Building a Sustainable Commitment Across the Institution*, 2nd ed. Stylus, publisher's information at <http://bit.ly/1rhc2Ac>. Amazon.com information at <http://amzn.to/1rh6DZQ>. Note the searchable "Look Inside" feature. An expurgated Google book preview is online at <http://bit.ly/1wkRxa8>. Note especially "Making the Invisible Visible in Physics," pp. 123-129.

Matthews, M.R. ed. 2014. *International Handbook of Research in History, Philosophy and Science Teaching*. Springer, publisher's information at <http://bit.ly/1sQQs8I>, including a preview at <http://bit.ly/1wc5QAq>. Amazon.com information at <http://amzn.to/1ovyAlk>, note the searchable "Look Inside" feature. An expurgated Google book preview is online at <http://bit.ly/1wc5QAq>.

McCloskey, M, A. Caramazza, & B. Green. 1980. "Curvilinear Motion in the Absence of External Forces: Naïve Beliefs About the Motion of Objects," *Science* **210**, 1139. An abstract is online at <http://bit.ly/1zgSGUF>.

McCloskey, M. 1983a. "Naïve theories of motion," in *Mental Models*, D. Gentner & A.L Stevens, eds., Erlbaum, pp. 299-324. An ERIC abstract is online at <http://eric.ed.gov/?q=ED223417>.

McCloskey, M. 1983b. "Intuitive Physics," *Scientific Americ*an **248**(4): 122-130; online as a 643 kB pdf at <http://bit.ly/1vMgIEJ>.

McCloskey, M. A. Washburn, & L. Felch. 1983. "Intuitive physics: The straight down belief and its origin," *Journal of Experimental Psychology: Learning, Memory, and Cognition* **9**(4): 636-649; an abstract is online at <http://bit.ly/1wqo9Ai>.

McCutcheon, L.E. 1991. "A New Test Of Misconceptions About Psychology," *Psychological Reports* Volume 68: 647-653. The first page is online at <http://bit.ly/1oeWZvc>. The summary reads:

> "Some of the problems surrounding the use of true-false tests of psychological misconceptions are discussed. The development of a new 62-item, multiple-choice test of misconceptions designed to reduce these problems is outlined. The test was given to 79

students in introductory psychology. Reliability, validity, and normative data are provided. A comparison of the interest level of each item's topic allows psychology professors a reasonable criterion for deciding which misconceptions to discuss in class."

McDermott L.C. & E.F. Redish. 1999. "RL-PER1: Resource letter on physics education Research," *Am. J. Phys.* **67**(9): 755-767; online as a 127 kB pdf at <http://bit.ly/10hjC4R>.

McKinnon, J.W. 1971. "Earth science, density, and the college freshman," *J. Geol. Educ.* **19**(5): 218-220; an ERIC abstract, online at <http://eric.ed.gov/?id=EJ047690> reads:
"Reports the results of testing 143 college freshmen on the meaning of density. Relates reasons for student inability to conceptualize density to the hierarchy of experiences which leads to the understanding of the density concept."

Meltzer, D.E. & R. K. Thornton. 2012. "Resource Letter ALIP-1: Active-Learning Instruction in Physics," *Am. J. Phys.* **80**(6): 478-496, online to subscribers at <http://bit.ly/O35gtB>.

NRC. 2013. *Adapting to a Changing World - Challenges and Opportunities in Undergraduate Physics Education*. National Academy Press, online at <http://bit.ly/126os6j>.

Nuhfer, E. & D. Knipp. 2003. "The Knowledge Survey: A Tool for All Reasons," in *To Improve the Academy* **21**: 59-78; online as a 303 kB pdf at <http://bit.ly/1w9FUnx>.

Paden, D.W. & M.E. Moyer. 1969. "The Relative Effectiveness of Teaching Principles of Economics," *Journal of Economic Education* **1**(1): 33-45, online to subscribers at <http://bit.ly/1pUTplk>.

Poincaré, H. 1905. *Science and Hypothesis*, Walter Scott Publishing; online at <http://bit.ly/9hVfA8> thanks to the "Mead Project." A Wikipedia entry on Poincaré is at <<http://bit.ly/b4jGVS>>.

Redish, E.F. 1999. "Millikan lecture 1998: Building a science of teaching physics," *Am. J. Phys.* **67**(7): 562-573; online at <http://bit.ly/pePmuK>.

Schwartz, B.M. & R.A.R Gurung, 2012. *Evidence-Based Teaching for Higher Education*. American Psychological Association, publisher's information at <http://bit.ly/1sJ407w>. Amazon.com information at <http://amzn.to/1sz1qk8>.

Scriven, M. 2004. "Re: pre- post testing in assessment," AERA-D post of 15 Sept 2004 19:27:14-0400; online on the OPEN! AERA-D archives at <http://bit.ly/lPdUPw>.

Selingo, J. 2013b. *College Unbound: The Future of Higher Education and What It Means for Students*. New Harvest, publisher's information at <http://bit.ly/13vIgP0>. Amazon.com information at <http://amzn.to/17ZQ3pg>, note the searchable "Look Inside" feature.

Shadish, W.R., T.D. Cook, & D.T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin. Amazon.com information at

<http://amzn.to/JWWMVi>. A goldmine of references to the social-science literature of experimentation.

Shavelson, R.J. & L. Huang. 2003. "Responding Responsibly To the Frenzy to Assess Learning in Higher Education," *Change Magazine*, January/February; online as a 123 kB pdf at <http://bit.ly/1gg8WYA>.

Shavelson, R.J. 2008. "Formative Assessment," Guest Editor's Introduction, special issue of *Applied Measurement in Education*; online at <http://bit.ly/nn2Rcz>. Also free online at the same URL are five articles on formative assessment that appear in *Applied Measurement in Education* **21**(4) of 4 October; online to subscribers at <http://bit.ly/phtxnG>.

Simkins, S.P. & M.H. Maier. 2008. "Learning from Physics Education Research: Lessons for Economics Education," chapter 36 of *International Handbook on Teaching and Learning* [Hoyt & McGoldrick (2012); also online at the "Social Science Research Network" (SSRN) at <http://bit.ly/1rr962K> (click on "Download This Paper").

Suskie, L. 2004. "Re: pre- post testing in assessment," ASSESS post of 19 Aug 2004 08:19:53-0400; online on the OPEN! ASSESS archives at <http://bit.ly/9yIOyf >. Suskie makes 8 anti-pre/post-testing points to which Hake (2004) and Scriven (2004) respond.

Suskie, L. 2009. *Assessing Student Learning: A Common Sense Guide*, foreword by Trudy Banta. Wiley, 2nd edition, publisher's information at <http://amzn.to/fMXH97>, note the searchable "Look Inside" feature. Contrast this book by the anti-pre/post Suskie – see Suskie (2004) - with that of the pro-pre/post Peggy Maki (2011).

Foreword writer Trudy Banta appears to be as unaware of physics education research as is Suskie. Economist Bill Goffe (2011), in his PhysLrnR post "Re: Business agenda for K-12 STEM education: not research-informed" wrote: (paraphrasing):

> ". . . it appears that Physics Education Research isn't widely known even in higher ed. For example, Trudy Banta and Charles Blaich in a *Change Magazine* article "Closing the Assessment Loop" <http://bit.ly/lQyEYp> [Banta & Blaich (2011)] bemoan the fact that they can find very few instances of improved learning after a teaching innovation. The extensive physics education research that so convincingly demonstrates such a connection is not even mentioned."

Thornton, R.K. & D.R. Sokoloff. 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**(4): 338-352. An abstract is online at <http://bit.ly/ZGApjq>.

Vaughan, E.D. 1977. "Misconceptions about psychology among introductory psychology students," *Teaching of Psychology* **4**(3): 138-141. An abstract, online at <http://bit.ly/1rQBETt>, reads:

> "Common misconceptions about behavior are distressingly resistant to change by text reading and class discussion."

Wilson, M.R. & M.W. Bertenthal, eds. 2005. *Systems for State Science Assessment*, Nat. Acad.

Press; online at <http://bit.ly/f6WFeg>, click on "Download Free PDF."

Wittmann, M. 2007. "Re: Pretest and Post-test instruments for an Introduction to Psychology," PhysLrnR post of 22 Jan 2007 21:24:16-0500, online on the CLOSED PhysLrnR archives at <http://bit.ly/1sV4APA>. Post of 22 Jan 2007 21:24:16-0500 to PhysLrnR.

Wood, W.B., & J.M. Gentile. 2003. "Teaching in a research context," *Science* **302**: 1510; 28 November; online as a 213 kB pdf <http://bit.ly/SyhOvL>, thanks to Ecoplexity <http://bit.ly/152aFQ9>.