Bao *et al.*'s Comparison of Learning And Scientific Reasoning In Chinese & U.S. Schools: Alternate Conclusions and Recommendations †

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

ABSTRACT

The features, findings, conclusions, and recommendations of the valuable 2009 *Science* report "Learning and Scientific Reasoning: Comparisons of Chinese and U.S. students show that content knowledge and reasoning skills diverge" by Bao *et al*. are summarized.

The primary feature is that Chinese and U.S. students enrolled in introductory physics courses for science and engineering majors in medium-rated universities were tested near the start of classes with the *Force Concept Inventory* (FCI), the *Brief Electricity and Magnetism Assessment* (BEMA), and the *Lawson Classroom Test of Scientific Reasoning* (LCTSR).

The primary finding is that although Chinese students' averaged scores on the FCI and BEMA indicated *good conceptual understanding* of basic physics areas, and were, respectively, 1.9 and 2.9 standard deviations above those of U.S. students, their scores on the LCTSR were about the same as those of the U.S. students: at the low end of Lawson's hypothetical-deductive reasoning range.

Bao *et al*. draw the conclusion that "*current education and assessment in the STEM disciplines often emphasizes factual recall over deep understanding of science reasoning*" and recommend that researchers and educators (1) "invest more in the development of a balanced method of education, such as incorporating more *inquiry-based learning*," and (2) measure "not only content knowledge but also other factors so as to obtain a more holistic evaluation of students."

I criticize the Bao *et al*. report on two counts: (1) the conclusion doesn't follow from the findings; and (2) the recommendations are either ambiguous, problematic, or invalid.

A conclusion consistent with the findings is: *Average scores by both Chinese and U.S. freshmen on FCI, BEMA, and LCTSR indicate that K-12 STEM education in both those countries emphasizes factual recall over conceptual understanding and scientific reasoning*.

Recommendations that are, in my view, less ambiguous, less problematic, and more valid are: K-12 educators should (1) utilize interactive engagement, inquiry, cognitive enhancement methods, and tests of reasoning; (2) emphasize a few fundamental concepts of STEM; and (3) develop age-appropriate assessments of concepts, epistemological beliefs, learning attitudes, and reasoning so as to *formatively* assess the effectiveness of their teaching methods. In addition, *in the U.S.* efforts should be made to (4) reduce poverty, (5) upgrade the education, salary, and prestige of K-12 teachers, and (6) establish National Education Standards.

---

† The reference is: Hake, R.R. 2012. "Bao *et al*.'s Comparison of Learning And Scientific Reasoning In Chinese & U.S. Schools: Alternate Conclusions and Recommendations," online as ref. 66 at <http://bit.ly/a6M5y0>.

 Partially supported by NSF Grant DUE/MDR-9253965.

TABLE OF CONTENTS

I. Summary of the Features, Findings, Conclusions, and Recommendations of the *Science* Article by Bao *et al.* (2009a):

A. Features

1. The 1995 version [Halloun *et al*. (1995)] of the *Force Concept Inventory* [Hestenes *et al*. (1992)]; the *Brief Electricity and Magnetism Assessment* (BEMA) [Ding et al. (2006)]; and the 24-question version [in the Appendix of Coletta & Phillips (2005)] of the Lawson Classroom Test of Scientific Reasoning (LCTSR) [Lawson (1978)] were administered, before any college-level instruction was provided on the related content topics, to Chinese and U.S. students enrolled in introductory physics courses for science and engineering majors in medium-rated universities.

2. According to Bao *et al*. (2009a,b), the Chinese students had taken algebra-based physics for 5 years in grades 8-12, purportedly emphasizing the development of conceptual understanding and problem solving, but presumably of the "traditional" type: passive-student lectures, algorithmic-problem homework & exams, and recipe labs.

3. Although U.S. students study physics-related topics within other general science courses only one third of high-school students enrolls in a *one year* physics course [Hehn & Neuschatz (2006)], in sharp contrast to the *5 years* of physics courses taken in grades 8-12 by *all* Chinese students. Of course, the proportion of U.S. students studied by Bao *et al*. who took one year of physics is doubtless higher than one third because those students were enrolled in a science and engineering major courses. On the other hand, since "interactive-engagement" * physics courses such as ASU's Modeling program <http://modeling.asu.edu/>; are so rare in the U.S., most of the high-school physics courses taken by the U.S. students studied by Bao *et al*. were probably of the "traditional" type that contributed little to students' conceptual understanding.

---

* "Interactive Engagement" (IE) courses are *operationally* defined [Hake (1998a)] as:

> "those designed at least in part to promote conceptual understanding through the active engagement of students in heads-on (always) and hands-on (usually) activities that yield immediate feedback through discussion with peers and/or instructors."

Thus a hallmark of IE courses is their use of *formative assessment* in the sense used by Black & Wiliam (1998) and Shavelson (2008):

> "All those activities undertaken by teachers -- and by their students in assessing themselves -- *that provide information to be used as feedback to modify teaching and learning activities*." [My *italics*.]

Thus IE courses are *not* the same as "inquiry-based" courses as that term is commonly understood – see Eq. (4) on page 12.]

B. Findings

1. Chinese students' average scores of 86% ± 14% on the FCI, and 66 ± 13% on the BEMA, indicate relatively *good* average conceptual understanding of Newtonian mechanics and E & M. Table III of Bao *et al*. (2009b) indicates that the Chinese students' average scores on the FCI and BEMA tests were, respectively 1.9 and 2.9 standard deviations above the U.S. students. Nevertheless, the "scientific reasoning" of the Chinese students as gauged by the LCTSR scores, 75% ± 16%, was on the low end of Lawson's "hypothetical-deductive" thinking range. . . . [[assuming the association of scores with thinking ranges for the 24-question version of the LCTSR scales with the 12-question version as indicated in the reference to Lawson (1995)]]. . . ., less than might be expected for science and engineering majors.

2. U.S students average scores of 49% ± 19% on the FCI, and 27 ± 10% on the BEMA, indicate relatively *poor* average conceptual understanding of Newtonian mechanics and E & M. Their "scientific reasoning" as gauged by the LCTSR scores, 74% ± 18%, was on the low end of Lawson's (1995) "hypothetical-deductive" thinking range, just as for the Chinese students.

C. Conclusion

From findings "B-1,2" above, Bao *et al*. (2009a) conclude that:

> "The results from this study are consistent with existing research [Schoenfeld (1988), Elby (1999), Linn *et al*. (2006), Zheng *et al*. (2008)] which suggests that *current education and assessment in the STEM disciplines often emphasizes factual recall over deep understanding of science reasoning*."

D. Recommendations

1. "Because students ideally need to develop both content knowledge and transferable reasoning skills, researchers and educators must invest more in the development of a balanced method of education, such as incorporating more *inquiry-based learning* [Zimmerman (2007), Adey & Shayer (1990), Lawson (1995), Benford & Lawson (2001), Marek & Cavallo (1997), and Gerber *et al*. (2001)] that targets both goals."

2. "As much as we are concerned about the weak performance of American students in TIMSS <http://nces.ed.gov/timss/> and PISA <http://www.pisa.oecd.org/>, it is valuable to inspect the assessment outcome from multiple perspectives. With measurements on not only content knowledge but also other factors, one can obtain a more holistic evaluation of students, who are indeed complex individuals."

II. Criticism #1: The Conclusion of Bao *et al.* Doesn't Follow From Their Findings

Conclusion "IC" above is, to repeat:

> "current education and assessment in the STEM disciplines often emphasizes *factual recall* over deep understanding of science reasoning" . . . . (IC)

does NOT follow from findings "IB-1,2" above (paraphrasing):

> "Chinese students' average scores on the FCI and BEMA indicate relatively *good* average conceptual understanding of Newtonian mechanics and E & M, but their average scores on the LCTSR indicate that their 'scientific thinking?' was on the low end of Lawson's 'hypothetical-deductive' thinking range." . . . . (IB-1)

> "U.S students average scores on the FCI and BEMA indicate relatively *poor* average conceptual understanding of Newtonian mechanics and E & M. Their average scores on the LCTSR indicate that their 'scientific reasoning' was on the low end of Lawson's 'hypothetical-deductive' thinking range". . . . .(IB-2)

A. Scores on *Concept Inventories* Gauge Students' *Conceptual Understanding*, NOT *Factual Recall*

The reason that conclusion (IC) doesn't follow from findings (IB-1,2) is that scores on the FCI and BEMA, as is the case for other *Concept Inventories* <http://en.wikipedia.org/wiki/Concept_inventory>, are ***not*** indicators of the degree of students' *factual recall*. Instead they are indicators of the degree of students' *conceptual understanding* (unless students are simply told the answers to questions similar to those on Concept Inventories or have somehow gained access to the Concept Inventories.

Concept Inventories such at the FCI have been developed through arduous qualitative and quantitative research by disciplinary experts - see e.g., (a) "Evidence on Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education: Workshop on Linking Evidence and Promising Practices in STEM Undergraduate Education" (National Academies, 2008); and (b) "The Impact of Concept Inventories On Physics Education and It's Relevance For Engineering Education" (Hake, 2011b).

An example of an FCI-like question [taken from Hake (2002a)] that probes for *conceptual understanding* of Newtonian mechanics is:

> A student in a lab holds a brick of weight *W* in her outstretched horizontal palm and lifts the brick vertically upward at a constant speed. While the brick is moving vertically upward at a constant speed, the magnitude of the force on the brick by the student's hand is:
>
> A. constant in time and zero.
> B. constant in time, greater than zero, but less than *W*.
> C. constant in time and *W*.
> D. constant in time and greater than *W*.
> E. decreasing in time but always greater than *W*.

Note that the responses include as distractors not only "D," the common Aristotelian misconception that "motion requires a net force," but also other, less common student misconceptions, "A" and "E," that might not be known to teachers lacking "pedagogical content knowledge" [Shulman (1986, 1987)]. The present distractors are based on my years of listening to students as they worked through the experiments in *Socratic Dialogue Inducing Lab #1* "Newton's First and Third Laws" (Hake 2001). For *actual* FCI questions, the distractors were usually gleaned through careful *qualitative* research involving interviews with students and the analysis of their oral and written responses to mechanics questions.

Judging from the results of my survey "Interactive-engagement vs traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses" [Hake (1998a)], after a traditional "TRAPT" (TRansmission to A Passive Target") physics course, probably only about 50% of students would give the correct answer: "C. constant in time and *W*."

In sharp contrast to the above conceptually-oriented question, an example of a traditional exam question that probes for "factual recall" is:

Newton's Second Law is:
A. F = 0
B. F = mv
C. F = ma
D. F = constant
E. F = m (mv)

After a traditional "TRAPT" physics course, probably about 99% of students would parrot the correct answer: "C. F = ma." But 99% would probably not have the vaguest idea of the *operational* meanings of F, m, or a – see "Laws of Classical Motion: What's F? What's m? What's a?" [Weinstock (1961)]; " 'What's F? What's m? What's a?': A Non-Circular SDI-TST-Lab Treatment of Newton's Second Law" [Hake & Wakeland (1997)]; and "Helping Students to Think Like Scientists in Socratic Dialogue Inducing Labs" [Hake (2012b)].

The unfortunate conflation of "factual recall" and "conceptual understanding" in Bao *et al.* (2009a,b) has been propagated in the media - to the detriment of the public's understanding of educational issues - in e.g.: (a) Bao's NPR interview "Learning Facts vs Learning to Reason" at <http://bit.ly/9eYSdc>; (b) the Ohio State news release "Study: Learning Science Facts Doesn't Boost Science Reasoning" [Gorder (2009)]; (c) the *Inside Higher Ed* report "Blinding Them With Science" [Moltz (2009)]; (d) the *Science Daily* (2009) report "College Freshmen In US And China: Chinese Students Know More Science Facts But Neither Group Especially Skilled In Reasoning"; and (e) "Explaining the Reasoning-Fact Gap" [Carlson (2009)].

Thus, in my opinion conclusion "IC" should have been phrased something like:

"current education and assessment in the STEM disciplines often emphasizes CONCEPTUAL UNDERSTANDING over deep understanding of science reasoning" . . . (IC*)

Although conclusion "IC*" now follows logically from findings "IB-1,2" it is nevertheless *problematic* on at least two counts:

1. As Paul Gross (2009) put it:

   "Although [Bao *et al*. (2009a)] refer to the physics courses, especially those taken by the Chinese students, as emphasizing 'conceptual physics understanding and problem-solving skills,' the researchers do not, apparently, include conceptual understanding and problem-solving skills within scientific reasoning ability. For this old subscriber to *Science*, such an exclusion is incomprehensible."

   Bao *et al*.'s exclusion of "conceptual understanding and problem-solving skill" from "scientific reasoning" is also, I suspect, incomprehensible to most physicists if (a) "conceptual understanding" means, e.g., understanding of Newtonian mechanics as gauged, for example, by scores on tests of conceptual understanding such as the FCI; and if (b) "problem-solving skill" means, e.g., ability to solve non-algorithmic problems, as gauged, for example, by scores on the *Mechanics Baseline* test of Hestenes & Wells (1992)].

2. It is probable that *neither* Chinese nor U.S. K-12 physics courses emphasize conceptual understanding:

   a. In U.S. schools, physics *per se* is not generally taught in the K-8 grades. Physics *may* be taught in grades 9-12, but such courses are usually of the "traditional" type and do not emphasize conceptual understanding – I suspect that "interactive engagement" courses such as ASU's Modeling program <http://modeling.asu.edu/>, which *do* emphasize conceptual understanding, constitute only a very small percentage of grade 9-12 physics courses in the U.S.

   b. Despite what might be concluded from the relatively high FCI scores reported by Bao *et al*. (2009a,b) for the Chinese freshmen in their study, it's probable that Chinese schools generally do *not* emphasize conceptual understanding. This is suggested by the fact that after taking physics for ***five years***, the average FCI score for Chinese students studied by Bao *et al*. was 86%. As shown below, the importance of the factor 5 longer physics "soak time" for Chinese as opposed to U.S. students - see e.g., "Soak Time [was Academically Adrift?]" [Hake (2011a)] – is vitally important in assessing the results of Bao *et al*., a point previously made by PhysLrnR's Bill Goeff (2009) and John (Texas) Clement (2009).

B. Relatively High Scores on Concept Inventories by Chinese Freshmen Do Not Mean That Chinese K-12 Education Emphasizes Conceptual Understanding

1. Assume that the average FCI normalized gain <g> for a course with conventional physics instruction in China in <g> = 0.28 [assumed slightly higher than the average <g> = 0.26 for 4 "traditional" U.S. high-school courses given in Table 1a of Hake (1998b) because of the selective character of (a) Chinese education, and (b) the sampled population - science and engineering majors enrolled in calculus-based introductory physics courses. An FCI normalized gain of <g> less than 0.30 places grade 8-12 Chinese physics courses in the “low gain” region in comparison with the 61 courses surveyed by Hake (1998a,b) – see Fig. 1 of Hake (1998a).

Here, as indicated previously the average normalized gain <g> is defined [Hake (1998a,b; 2012a) as

$$\langle g\rangle = (\langle\%post\rangle - \langle\%pre\rangle) / (100 - \langle\%pre\rangle) \quad \ldots\ldots\ldots\ldots\ldots (1)$$

2. Assume that for the FCI on the first year (grade 8) of Chinese physics instruction the class average percentage pretest score $\langle\%pre\rangle_1 = 30\%$ (10 % above the random guessing score and the same as the average <%pre> for high-school courses listed in Table 1a of Hake (1998b), where the angle brackets <...> indicate class averages.

3. Assume that <g> remains at 0.28 for all 5 years of conventional instruction. Then it's easy to show by successive iterations for years i = 1, 2, 3, 4, 5 of the equation:

$$\langle\%post\rangle_i = \langle\%pre\rangle_i + \langle g\rangle\, [100\% - \langle\%pre\rangle_i] \quad \ldots\ldots\ldots\ldots\ldots (2)$$

that at the end of the fifth year of instruction the posttest score would be 86%, the same as given by Bao *et al*. (2009a) in their table on p. 586 as the average pretest score of freshman Chinese students on the FCI.

Thus, even though Chinese physics instruction in grades 8-12 is by most accounts similar to “traditional” relatively ineffective (i.e., relatively low average-normalized-gain <g>) physics instruction in the U.S., *the factor 5 longer soak time results in relatively good performance on the FCI at the end of the 5th year*.

4. Thus, to reiterate: the conclusion of Bao *et al*. (2009a,b) that:

> “current education and assessment in the STEM disciplines often emphasizes ***factual*** *recall* over deep understanding of science reasoning” . . . . (IC)

even when corrected to:

> “current education and assessment in the STEM disciplines often emphasizes ***conceptual understanding*** over deep understanding of science reasoning”. . . .(IC*)

so as to be consistent with the reported Concept Inventory data, is still *problematic* because

(a) it’s not clear that conceptual understanding should not be included within scientific reasoning ability, as pointed out by Paul Gross (2009), and

(b) it’s likely that neither U.S. or Chinese grade 8-12 physics education emphasizes conceptual understanding.

III. Criticism #2: Recommendation "ID-1" Is Ambiguous

Recommendation "ID-1" above is (paraphrasing):

"To promote both *content knowledge* and *transferable reasoning skills*, a balanced method of education incorporating more *inquiry-based learning* [Zimmerman (2007), Adey & Shayer (1990), Lawson (1995), Benford & Lawson (2001), Marek & Cavallo (1997), and Gerber *et al*. (2001)] that targets both goals should be developed." . . . . (ID-1)

A. What Do Bao *et al.* (2009a) Mean By "Content Knowledge"?

Is it "figurative (or declarative) knowledge," "operative (or procedural) knowledge," both, or neither? The late Arnold Arons (1983) wrote [bracketed by lines "AAAA. . . ."; my insert at ". . . .[[insert]]. . . . ."; my *italics*]:

AAAAAAAAAAAAAAAAAAAAAAAA

Researchers in cognitive development describe two principal classes of knowledge. *Figurative (or declarative)* and *operative (or procedural)* [Anderson (1980); Lawson (1982)]. . . [[Caution – in the mathematics education literature see e.g., Kilpatrick *et al*. (2001) "procedural knowledge" means, e.g., knowledge of a mathematical *procedures* such as "invert and multiply" to compute [(a/b)/(c/d)] without necessarily knowing the rationale for the operation, thus nearly the opposite of what "procedural knowledge" means to cognitive researchers – see below]]. . . . Declarative knowledge consists of knowing "facts"; for example that the moon shines by reflected sunshine, that the earth and planets revolve around the sun, that matter is composed of discrete atoms and molecules; that animals breathe in oxygen and expel carbon dioxide. Operative knowledge, on the other hand, involves understanding the source of such declarative knowledge (How do we know the moon shines by reflected sunlight? Why do we believe the earth and planets revolve around the sun when appearances suggest that everything revolves around the earth? What is the evidence that the structure of matter is discrete rather than continuous? What do we mean by the names "oxygen" and "carbon dioxide" and how do we recognize these as different substances?) and *the capacity to use, apply, transform, or recognize the relevance of declarative knowledge in new or unfamiliar situations*.

AAAAAAAAAAAAAAAAAAAAAAAA

The above *italicized* sentence indicates that Arons regarded "operative knowledge" as including "transferable reasoning skills." But recommendation "ID-1" seems to imply that: "content knowledge" does *not* include "transferable reasoning skills." So I *suspect* (please correct me if my suspicion is wrong) that for Bao *et al*. (2009a,b):

"Content Knowledge" = "declarative knowledge" = "factual knowledge". . . . (3)

Paul Gross (2009) came to a similar conclusion regarding the use of "content knowledge" by Bao *et al*.. Gross wrote:

"[To Bao et al. (2009a)] scientific reasoning is not simply subsumed under content; the authors' use of 'content' implies that, for them, the word means something like just the facts, ma'am—with perhaps some very ad hoc concept juggling and problem solving."

Furthermore:

a. "Interactive Engagement" (IE) courses which emphasize conceptual understanding often stress "transferable reasoning skills" - see e.g.: (a) "Learning to Think Like Scientists with the PET Curriculum" [Otero & Gray (2007)]; (b) "Are Most People Too Dumb for Physics?" [Lasry et al. (2009)]; and (c) "Helping Students to Think Like Scientists in Socratic Dialogue Inducing Labs" [Hake (2012b)].

b. At least some IE courses DO enhance scientific reasoning as measured by the LCTSR - e.g.: "Addressing Barriers to Conceptual Understanding in IE Physics Classes" (Coletta & Phillips, 2009)] for university classes; "Influence of Three Different Methods of Teaching Physics on the Gain in Students' Development of Reasoning" (Marusic & Slisko, 2012) for a high-school class.

Lasry *et al*. (2009) wrote (see their article for the detailed references):

"In physics, there is a long tradition of modeling and explicitly teaching problem-solving skills. Students are encouraged not only to solve problems but reflect on the process. This thinking about one's thinking, or metacognition, whereby one monitors and reflects upon his or her thinking, is extremely helpful in learning science and mathematics. Furthermore, metacognition is also a common characteristic of expertise across domains [Bransford *et al*. (2000)]. Thus, reflecting on the thinking involved in context-rich problem solving develops metacognitive skills, which is one of the hallmarks of experts . . . .[[and scientific thinking]]. . . , regardless of their domain."

B. What Do Bao *et al.* (2009a) Mean By "Inquiry-Based Learning"?

In my opinion, the term *inquiry-based learning*, is, without an *operational definition*, essentially meaningless. The same can be said of other educational catch phrases, e.g.: *active learning, cooperative learning*, *direct instruction*, *discovery learning*, *hands-on activities*, etc. Paraphrasing Klahr & Li (2005):

"those engaged in discussions about implications and applications of educational research should focus on clearly defined instructional methods and procedures, rather than vague labels and outmoded *–isms*."

See also: (a) "Language Ambiguities in Education Research" [Hake (2008)]; (b) "Education Research Employing Operational Definitions Can Enhance the Teaching Art" [Hake (2010a)]; and (c) the attempt by Minner *et al*. (2009) to operationalize the term "inquiry-based science instruction" in "Inquiry-based science instruction—what is it and does it matter?" as discussed in "Re: Metastudy on impact of inquiry in K-12" [Hake (2010b)].

Because (a) Bao et al. (2009a,b) do not give an explicit definition of "inquiry-based learning," and (b) there are numerous and conflicting definitions of "inquiry-based learning" – se e.g., "Reforming Science Teaching: What Research Says About Inquiry" [Anderson (2002)] - one must assume that the meaning intended by Bao et al. is reflected in the references given by them for that term: Zimmerman (2007); Adey & Shayer (1990); Lawson (1995); Benford & Lawson (2001); Marek & Cavallo (1997); and Gerber, Cavallo, & Marek (2001). Considering each of these references:

a. Zimmerman's (2007) review focusses "on the reasoning and thinking skills involved in students' scientific inquiry, such as hypothesis generation, experimental design, evidence evaluation and drawing inferences."

b. I could not locate Bao *et al*.'s reference to "Adey & Shayer (1990)" – it may be a typo for "Adey & Shayer (1994) – but Shayer & Adey (2002) on pp. 4-6 of *Learning Intelligence: Cognitive Acceleration across the Curriculum from 5 to 15 Years* imply that "Cognitive Acceleration" attempts to set up environments "likely to provide maximal stimulation to the intellect" as suggested by the developmental psychology of Jean Piaget and the socio-cultural psychology of Lev Vygotsky, drawing from those theories six working principles which they refer to as the "pillars" of cognitive acceleration: schema theory, concrete preparation, cognitive conflict, social construction, metacognition, and bridging. Thus it would appear that "cognitive acceleration," just as "interactive engagement," is *not* synonymous with "inquiry learning" as the term is commonly understood – see Eq. (4) on page 12.

c. Lawson (1995) considers "inquiry" on pages 155-158. He writes:
"John Dewey was a vocal advocate of science instruction that emphasized a method of inquiry. Addressing the National Education Association, Dewey (1916) argued that

> 'science is primarily the method of intelligence at work in observation, in inquiry and experimental testing; that, fundamentally, what science means and stands for is simply the best ways yet found out by which human intelligence can do the work it should do, ways that are continuously improved by the very process of use'.

But it would take more than forty years before this view of the nature of science would make its way into a large-scale curriculum movement, or not until the National Science Foundation sponsored curriculum development projects in the late 1950s and early 1960s. . . . . . .A survey of various teaching methods advocated in science methods text books, just before and during this curriculum movement echoed the emphasis on inquiry. One textbook even included a method called the 'Learning Cycle' [Heiss et al. (1950). . . . . The Heiss *et al*. learning cycle most closely approximates a hypothetical-deductive learning cycle: It involves generating hypotheses and their testing through the deduction of their consequences."

d. Benford & Lawson (2001) on page 17 wrote [see that report for the references]:
"Inquiry Teaching – to quantify the effectiveness with which a teaching assistant used inquiry in the classroom, the *Reformed Teaching Observation Protocol* (RTOP) [Sawada et al. (2000)] was used. . . . [it quantifies the extent to which science and math teachers use inquiry techniques as defined by the National Research Council (1990, 1995), and American Association for the Advancement of Science (1990, 1993), and the National Council for the Teaching of Mathematics (1989, 1991, 1995)"

I could not locate the Benford/Lawson references NRC (1990, 1995) but *National Science Education Standards* [NRC (1996)] states:

"Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations. Students will engage in selected aspects of inquiry as they learn the scientific way of knowing the natural world, but they also should develop the capacity to conduct complete inquiries" . . . .(4)

The above definition of "inquiry" appears to be consistent with the use of the term in:

(a) NRC's (2011b) *Framework for K-12 Science Education* wherein it is stated [my *italics*]:
" . . . *because the term 'inquiry,'* extensively referred to in previous standards documents, *has been interpreted over time in many different ways throughout the science education community*, part of our intent in articulating the practices in Dimension 1 . . . .[["Practices" Chapter 3]]. . . is to better specify what is meant by inquiry in science and the range of cognitive, social, and physical practices that it requires. As in all inquiry-based approaches to science teaching, our expectation is that students will themselves engage in the practices and not merely learn about them secondhand. Students cannot comprehend scientific practices, nor fully appreciate the nature of scientific knowledge itself, without directly experiencing those practices for themselves." On page 3-5 it is stated: We consider eight practices to be essential elements of the K-12 science and engineering curriculum:

(1) Asking questions (for science) and defining problems (for engineering)
(2) Developing and using models
(3) Planning and carrying out investigations
(4) Analyzing and interpreting data
(5) Using mathematics, information and computer technology, and computational thinking
(6) Constructing explanations (for science) and solutions (for engineering)
(7) Engaging in argument from evidence
(8) Obtaining, evaluating, and communicating information

(b) McDermott *et al.*'s (2005)] *APS Forum on Education* article "Physics by Inquiry." They wrote:
"In all of the modules in *Physics by Inquiry* (PbI). . . . .[[McDermott and Physics Education Group (University of Washington) (1996)]]. . . . there is a strong emphasis on the development of important scientific skills, such as distinguishing between observations and inferences, controlling variables, proportional reasoning, deductive and inductive reasoning, etc. PbI fosters the simultaneous development of physical concepts, reasoning ability, and representational skills within a coherent body of content. The teachers go through the reasoning in depth and are guided in synthesizing what they have learned into a coherent conceptual framework. Since effective use of a particular instructional strategy is often content-specific, instructional methods are taught by example. If teaching methods are not studied in the context in which they are to be implemented, teachers may be unable to identify the elements that are critical. Thus they may not be able to adapt an instructional strategy that has been presented in general terms to specific subject matter or to new situations."

e. Marek & Cavallo (1997) in "The Learning Cycle: Elementary School Science and Beyond" advocate (according to the publisher's information at <http://bit.ly/zH6aDv>)

> "more than a classroom strategy; it is a philosophy of education - a model of instruction that can promote critical thinking and meaningful learning. It places students at the center of their learning experiences, encouraging them to engage in explorations, form new understandings, and relate those understandings to other concepts."

f. Gerber, Cavallo, & Marek's (2001) "Relationships among informal learning, environments, teaching procedures, and scientific reasoning ability," is not available to me, but it's probably safe to assume that their take on "inquiry learning" is very similar to that of Marek & Cavallo (1997) above.

After reviewing "a" – "f" above I *suspect* (please correct me if I'm wrong) that for Bao *et al*. (2009a,b):

"Inquiry-Based Learning" means essentially the same as is meant in NRC (1996) – Eq. (4) above - plus Shayer & Adey's "cognitive acceleration". . . . .(5)

Bao et al. (2009a,b) imply that "inquiry-based learning" will enhance students' scientific reasoning abilities; for example Bao *et al*.'s (2009a) recommendation I-D1 reads:

"To promote both *content knowledge* and *transferable reasoning skills*, a balanced method of education incorporating more *inquiry-based learning* . . . . .that targets both goals should be developed." . . . .(ID-1)

But as far as I know (please correct me if I'm wrong):

a. There have been *only two reports* (both by ASU groups) containing evidence that inquiry courses improve students scientific reasoning as gauged by the LCTSR: "Relationships Between Effective Inquiry Use and the Development of Scientific Reasoning Skills in College Biology Labs" (Benford & Lawson, 2001), and "Changing the Culture of Undergraduate Science Teaching" (Wyckoff, 2001).

b. There have been *no reports* containing evidence that "cognitive acceleration" courses improve students scientific reasoning as gauged by the LCTSR.

c. The extent to which inquiry methods promote conceptual understanding as gauged by "Concept Inventories" has not been reported. It's not easy to see how "inquiry methods" as set forth in Eq. (4) could, for example, promote students' crossover from the common-sense Aristotelian World of 'motion requires a force" to the abstract counter-intuitive Newtonian World – see "Promoting Student Crossover to the Newtonian World" [Hake (1987)].

IV. Criticism #3: Recommendation "ID-2" Is Invalid, Problematic, and Ambiguous

Recommendation "ID-2" above is (paraphrasing):

> "As much as we are concerned about the weak performance of American students in TIMSS <http://nces.ed.gov/timss/> and PISA <http://www.pisa.oecd.org/>, it is valuable to inspect the assessment outcome from multiple perspectives. With measurements on not only *content knowledge* but also *other factors*, one can obtain a more holistic evaluation of students.". . . . .(ID-2)

The substitution of *factual knowledge* for *content knowledge* in accord with Eq. (3) on page 9 leads to:

> "As much as we are concerned about the weak performance of American students in TIMSS <http://nces.ed.gov/timss/> and PISA <http://www.pisa.oecd.org/>, it is valuable to inspect the assessment outcome from multiple perspectives. With measurements on not only *factual knowledge* but also other factors, one can obtain a more holistic evaluation of students.?". . . . . ((ID-2*)

Although ID-2* is now phrased in language consistent with the rest of of Bao et al., it is:

A. Invalid Because Neither TIMMS Nor PISA Measures Outcomes of *ONLY Factual Knowledge*

According to the "Comparing Features of the Assessments" NCES (2008?): "TIMSS identifies knowing, applying and reasoning as its three cognitive categories . . . . TIMSS includes an overarching dimension called scientific inquiry, which attempts to measure students' abilities to engage in (paper-and-pencil) inquiry tasks. . . . . . The PISA science literacy framework also has content and cognitive dimensions. . . . . PISA's content dimension includes both knowledge of the natural world (in the fields of life systems, physical systems, Earth and space systems, and technology systems) and knowledge about science itself (scientific inquiry and scientific explanations). PISA's cognitive dimension describes important competencies required for scientific literacy: identifying scientific issues, explaining scientific phenomena, and using scientific evidence. In the PISA framework model, the competencies are prominent—they form the subscales for reporting,. . . ."

B. Problematic Because It's Not Clear That *Factual Knowledge* Can Be Separated From *Reasoning*.

Paul Gross (2009) in "Learning Science: Content - With Reason" wrote:

> "At least among cognitive scientists, the consensus seems to be that, 'Just as it makes no sense to try to teach factual content without giving students opportunities to practice using it, it also makes no sense to try to teach critical thinking devoid of factual content.'. . . .[["Critical Thinking: Why Is It So Hard to Teach?" (Willingham, 2007a); see also "How Knowledge Helps: It Speeds and Strengthens Reading Comprehension, Learning - and Thinking" (Willingham, 2007b)]]. . . . . Here, for 'critical thinking,' we may substitute 'scientific reasoning.' In the relevant contexts, they mean almost the same thing: scientific reasoning in the absence of scientific content doesn't make sense. Reasoning and content are not practically and neatly separable."

Shavelson & Huang (2003) made a similar point in "Responding Responsibly To the Frenzy to Assess Learning in Higher Education." They wrote (p.13):

> ". . . we know from research that learning, at least initially, is highly situated and context dependent. Only through extensive engagement, practice, and feedback within a particular subject area does learned knowledge become sufficiently decontextualized to enable it to transfer to the realm of enhanced reasoning, problem-solving, and decision-making skills exercises in broader or multiple domains.

C. Ambiguous Because *Other Factors* Is Undefined

It seems likely that among the "other factors" that Bao et al. had in mind was "scientific reasoning" as gauged by *Lawson Classroom Test of Scientific Reasoning* (LCTSR) designed to test the degree of students' hypothetical-deductive reasoning range. However there's been disagreement as to whether or not hypothetical-deductive reasoning is *the* hallmark of scientific reasoning:

(1) Douglas Allchin (2003) in "Lawson's Shoehorn, or Should the Philosophy of Science Be Rated 'X'? " claimed that Lawson's (2002) interpretation of Galileo's discovery of the moons of Jupiter as illustrating the role of hypothetical-deductive reasoning in science exhibits "several historical errors" and illustrates "how philosophical preconceptions can distort history and thus lessons about the nature of science." A debate ensued among Lawson (2003, 2004), Allchin (2004, 2006a,b), and Brush (2004). See also Wivagg & Allchin (2002).

(2) Physicist Helen Quinn (2009), in "What is science" wrote:

"Theories and models develop over time. Based on data, they undergo a long-term process of testing and refinement before becoming accepted scientific explanations or tools in a given domain. Contrast that with the usual description of the scientific method, which reduces continuous and iterative theory building to *the idea that one makes and tests hypotheses*.. . . [[my *italics*]]. . . . The use of a broad theoretical framework within which each hypothesis must fit, and that gets refined by each test, is generally lacking in the textbook account."

(3) Philosopher Susan Haake (2003) in *Defending Science* wrote [as quoted by Gross (2009)]:

"Scientific inquiry is continuous with the most ordinary of everyday empirical inquiry. There is no mode of inference, no 'scientific method,' exclusive to the sciences and guaranteed to produce true, more nearly true, or more empirically adequate results. . . . And, as far as [science] is a method, it is what historians or detectives or investigative journalists or the rest of us do when we really want to find something out: make an informed conjecture about the possible explanations of a puzzling phenomenon, check how it stands up to the best evidence we can get, and then use our judgment whether to accept it, more or less tentatively, or modify, refine, or replace it."

(4) Philip Adey (2010), in “Thinking in Science – Thinking in General?” wrote [see his article for references other than Demetriou et al. (1992) (now available in a 1994 paperback)]:

> “Intuitively . . . a total separation of different types of thinking does not seem plausible and in fact the psychological evidence is clear (Anderson, 1992; Carroll, 1993) that there is always a significant correlation between higher level thinking across all different subject domains. Notwithstanding claims for completely independent ‘multiple intelligences’, all of the evidence points to the existence of one general intellectual processing mechanism (general intelligence, or ‘g’), which is supplemented by a range of specialised abilities such as verbal, quantitative, and spatial (Demetriou, Gustafsson, Efklides, & Plastidou, 1992).”

(5) Albert Einstein (1936), in *Physics and Reality* wrote:

> “The whole of science is nothing more than a refinement of everyday thinking. It is for this reason that the critical thinking of the physicist cannot possibly be restricted to the examination of concepts from his own specific field. He cannot proceed without considering critically a much more difficult problem, the problem of analyzing the nature of everyday thinking.”

(6) Paul Gross (2009) in “Learning Science: Content - With Reason,” identifies “scientific reasoning” with “inquiry.” He wrote:

> “Scientific reasoning goes by different names, one of the most favored being "inquiry," as in "inquiry-based learning." This type of science study is so well established in the United States that a book-length *retrospective* and prospective account of inquiry-based science standards was published by the U.S. National Research Council nearly a decade ago [NRC (2000)]. . . . . . . The current Science Framework for the National Assessment of Educational Progress [NAGB (2009)] reflects that preoccupation by dividing attention between science content and science practices. Of the latter, there are four, each preceded by an action verb: “identifying” or “using.” The “using” statements are explicit reasoning skills. . . . .[More recently see the *Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* [NRC (2011b)] and reviews by Helen Quinn (2011) and Paul Gross (2011).]

But regardless of whether or not hypothetical-deductive reasoning is considered to be *the* hallmark of scientific reasoning, it's important that students develop reasoning skills, as well stated by Anton Lawson (2010) in his review of *College Teaching and the Development of Reasoning* [Fuller et al. (2009)]:

> ". . . . the problem initially identified by McKinnon and Renner (1971) remains. Far too many college freshmen enroll with poorly developed reasoning skills. And most college and university instructors are still unaware of the extent of the problem and its solution. So they continue to emphasize the transmission of information to the detriment of student intellectual development. Unfortunately, in the USA at least, most college instructors are not hired for their pedagogical awareness and expertise. Rather they are hired for their subject matter and research expertise. So they continue in their ignorant bliss to teach the wrong things in the wrong way. *And future teachers continue to teach as they have been taught*. . . . .[My *italics*.]. . . .Consequently, in spite of near universal agreement among pedagogical experts that a serious problem remains and that a pedagogical solution exists, no easy way has been found to implement that solution on the broad scale that is necessary. Nevertheless, a good start could be made by making the present book required reading for all college and university instructors."

V. A Conclusion Consistent With the Findings of Bao *et al.*

The average scores by both Chinese and U.S. freshmen on the *Force Concept Inventory* (FCI), the *Brief Electricity and Magnetism Assessment* (BEMA), and the *Lawson Classroom Test of Scientific Reasoning* (LCTSR) indicate that K-12 STEM disciplines in both China and the U.S. often emphasize factual recall over conceptual understanding and scientific reasoning.

VI. Recommendations to Enhance the Effectiveness of K-12 Education That Are Less Ambiguous and More Valid Than Those of Bao *et al.* ....

A. For K-12 Educators and Researchers

To promote not only factual recall, but also conceptual understanding, transferable reasoning abilities, and understanding of scientific method, K-12 educators and researchers should:

(1) Utilize:

(a) "interactive engagement" ["Lessons from the Physics Education Reform Effort" (Hake, 2002a];

(b) "inquiry" [*National Science Education Standards* (NRC, 1996)], *Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (NRC 2011b)];

(c) "cognitive acceleration" [*Learning Intelligence: Cognitive Acceleration across the Curriculum from 5 to 15 Years* (Shayer & Adey, 2002); "Re: Cognitive Acceleration" (Hake (2011c)]; and "cognitive modifiability" ["Instrumental Enrichment - An Intervention Program for CognitiveModifiability" (Feuerstein et al., 1980); "Feuerstein's Instrumental Enrichment"(Hake, 2011d); "Re: active learning needs a theory" (Hake, 2006b)];

(d) tests of reasoning (see "3d" below) to gauge the cognitive level of students.

(2) Emphasize a few fundamental concepts of STEM ["Towards Coherence in Science Instruction: A Framework for Science Literacy" (Schmidt et al., 2011)].

(3) Develop (where necessary) K-12 versions of the assessments that have been used primarily for undergraduates - shown below within square brackets [...]:

(a) concepts ["Evidence on Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education" (National Academies, 2008)];

(b) epistemological beliefs ["Epistemological beliefs assessment for physical science" (Elby *et al*., undated];

(c) learning attitudes ["New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey" (Adams *et al*., 2006)];

(d) reasoning tests, e.g.: ["Lawson Classroom Test of Scientific Reasoning (Lawson, 1978); "Critical thinking Assessment Test" (CAT, 2011); "Collegiate Learning Assessment" (CAE, 2012); "Watson-Glaser Critical Thinking Appraisal" (COD, 2012); "The California Critical Thinking Skills Test (Facione, 1990); "Health Sciences Reasoning Test" (IA (2012);

all the above assessments "a" – "d" so as to *formatively* assess the effectiveness of teaching methods – not to be confused with the counter productive *summative* assessment mandated by the "No Child Left Behind" (NCLB) act – see e.g. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education* [Ravitch (2011)]. Here "formative" is used in the "Joint Committee on Standards for Educational Evaluation" [JCSEE (1994)] sense: "Formative evaluation is evaluation designed and used to improve an object, especially when it is still being developed."

B. For Those Concerned With K-12 Education in the U.S.

(4) reduce poverty ["The Overriding Influence of Poverty on Children's Educational Achievement" (Hake, 2011e, and references therein)];

(5) upgrade the education, salary, and prestige of K-12 Teachers ["The Need For Improved Physics Education of Teachers" (Hake, 2000a); "The General Population's Ignorance of Science Related Societal Issues: A Challenge for the University" (Hake, 2000b); "Whence Do We Get the Teachers - Response to Madison" (Hake, 2002b); *A Good Teacher in Every Classroom: Preparing the Highly Qualified Teachers Our Children Deserve* (Darling-Hammond, 2005); *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do* (Darling-Hammond & Bransford, 2005); "Diane Ravitch shares her views on education reform" (Gutierrez, 2012)]; and

(6) establish National Education Standards ["National Education Standards for the United States?" (Hake, 2009); *International Lessons About National Standards* (Schmidt, Houang, & Shakrani, 2009); "Soaring Systems: High Flyers All Have Equitable Funding, Shared Curriculum, and Quality Teaching" (Darling-Hammond, 2010-2011)].

References [All URL's accessed on 2 Feb 2012; most shortened by <http://bit.ly/>. Discussion list posts on the CLOSED! **: – (** PhysLrnR archives can be accessed 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!]

Adams, W.K., K.K. Perkins, N.S. Podolefsky, M. Dubson, N.D. Finkelstein, and C.E. Wieman. 2006. "New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey," *Phys. Rev. ST Phys. Educ. Res.* **2**, 010101; online as a 500 kB pdf at <http://bit.ly/xL1xU3>.

Adey, P. & M. Shayer. 1990. *J. Res. Sci. Teach.* **27**: 267; I failed to find the title of this article or any online indication of its existence – "1990" may be a typo for "1994", see below; but see Adey (2010), Adey & Shayer (1994), and Shayer & Adey (2002).

Adey, P. & M. Shayer, eds. 1994. *Really Raising Standards: Cognitive Interventions and Academic Achievement*. Routledge. Amazon.com information at <http://amzn.to/cLItGT>, note the searchable "Look Inside" feature. See also the more recent Shayer & Adey (2002). For references to the Adey/Shayer and related Feuerstein/Tribus literature see e.g., "Re: active learning needs a theory" [Hake (2006c)].

Adey, P., B. Csapo, A. Demetriou, J. Hautamaki, & M. Shayer. 2007. "Can we be intelligent about intelligence? Why education needs the concept of plastic general ability," *Educational Research Review* **2**: 75-97; online as a 418 kB pdf at <http://bit.ly/e4VatZ>.

Adey, P. 2010. "Thinking in Science – Thinking in General?" *Pantaneto Forum* **40**: October; online at <http://bit.ly/9h6kBg>. See also Adey et al. (2007).

Allchin, D. 2003. "Lawson's Shoehorn, or Should the Philosophy of Science Be Rated 'X'? " *Science & Education* **12**: 315–329; online as a 102 kB pdf at Allchin's site <http://bit.ly/g27BEE>. The abstract reads: "Lawson's (2002) interpretations of Galileo's discovery of the moons of Jupiter and other cases exhibit several historical errors, addressed here both specifically and generally. They illustrate how philosophical preconceptions can distort history and thus lessons about the nature of science." See the response by Lawson (2004a) and the ensuing debate among Allchin (2004, 2006a,b), Lawson (2004b), and Brush (2004).

Allchin, D. 2004. "Pseudohistory and Pseudoscience," *Science & Education* **13**: 179-195; online at Allchin's site <http://bit.ly/g27BEE>.

Allchin, D. 2006a. "Why Respect for History and Historical Error Matters," *Science & Education* **15**: 91-111; online at Allchin's site <http://bit.ly/g27BEE>.

Allchin, D. 2006b. "Lawson's Shoehorn, Reprise," *Science & Education* **15**:113-120; an abstract is online at <http://bit.ly/z6DBzB>.

Anderson, L.W. & L.A. Sosniak, eds. 1994. *Bloom's Taxonomy: A Forty-Year Retrospective*, Ninety-Third Yearbook of The National Society for the Study of Education, Univ. of Chicago Press. Amazon.com information at <http://amzn.to/wdatap>.

Anderson, L.W. & D. Krathwohl, eds. 2001. *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Addison Wesley Longman. See also Anderson & Sosniak (1994). The original 1956 Bloom et al. cognitive domain taxonomy has been updated to include important post-1956 advances in cognitive science - see especially Chapters 4 & 5 on the "Knowledge and Cognitive Process Dimensions." Amazon.com information at <http://amzn.to/wLN3yV>.

Anderson, J.R. 1980. *Cognitive Psychology and Its Implications*. Freeman, 1st ed. This popular text is now in its 7th edition [Anderson (2009)].

Anderson, J.R. 2009. *Cognitive Psychology and Its Implications*. Worth Publishers, 7th ed., publisher's information at <http://bit.ly/wRtZTs>. Amazon.com information at <http://amzn.to/wNAZOs>.

Anderson, R.D. 2002. "Reforming Science Teaching: What Research says about Inquiry," *Journal of Science Teacher Education* **13**(1): 1-12; online as a 139 kB pdf at <http://bit.ly/wwnZku>. Based on a commissioned paper prepared for the Center for Science, Mathematics and Engineering Education at the National Research Council. Anderson wrote:

> "Inquiry has a decades-long and persistent history as the central word used to characterize good science teaching and learning. Even at a time when a new word, *constructivism* . . . . [[see e.g., the Wikipedia entry at <http://bit.ly/zxj572>]]. . . . had entered the general educational lexicon as the descriptor of good education, the authors of the national science education standards (NSES) chose to stay with inquiry and totally ignore the new word. . . . [[For the pro and cons of 'constructivism' in education see 'Constructivist Instruction: Success or Failure?' (Tobias & Duffy, 2009)]]. . . .
>
> But in spite of its seemingly ubiquitous use, many questions surround inquiry. *What does it mean to teach science as, through, or with inquiry?* . . . . [[my *italics*]]. . . Is the emphasis on science as inquiry, learning as inquiry, teaching as inquiry or all of the above? Is it an approach to science education that can be realized in the classroom or is it an idealized approach that is more theoretical than practical? Is it something that the 'average' teacher can do, or is it only possible in the hands and minds of the exceptional teacher? What are the goals of its use? Does it result in greater or better learning? How does one prepare a teacher to utilize this type of science education? What barriers must be overcome to initiate such science education in the schools? What dilemmas do teachers face as they move to this form of science education? The list of questions goes on. They are of particular importance to people committed to the NSES and wanting to see these standards put into greater practice.
>
> Reformers from all categories—teachers, teacher educators, administrators, policy makers and members of the general public want to know what answers research has for such questions. Given the central role of teacher education in the process of educational reform, however, these questions are of particular interest to science teacher educators. Researchers' pursuit of answers has resulted in an extensive literature. Defining the arena broadly, the number of studies is in the hundreds and probably more. This body of research literature is worth exploring, but it will be necessary to limit and focus."

Arons, A.B. 1983. “Achieving Wider Scientific Literacy,” *Daedalus*, Spring. Reprinted in Arons (1997) as Chapter 12 “Achieving Wider Scientific Literacy.”

Arons, A.B. 1997. *Teaching Introductory Physics*. Wiley, publisher's information at <http://bit.ly/jBcyBU>. Amazon.com information at <http://amzn.to/bBPfop>, note the searchable “Look Inside” feature.

Bao, L., T. Cai, K. Koenig, K. Fang, J. Han, J. Wang, Q. Liu, L. Ding, L. Cui, Y. Luo, Y. Wang, L. Li, & N. Wu. 2009a. “Learning and Scientific Reasoning: Comparisons of Chinese and U.S. students show that content knowledge and reasoning skills diverge.” *Science* **323**(5914): 586 - 587, 30 January; online to subscribers at <http://bit.ly/dVR9Yx>, online to all as a 184 kB pdf at <http://bit.ly/90sdAG>; supporting material is online as a 152 kB pdf at <http://bit.ly/aNmbVz>. See also the longer report “Learning of content knowledge and development of scientific reasoning ability: A cross culture comparison” [Bao et al. (2009b)] in the *American Journal of Physics*. For the relevance of the findings of Bao et al. (2009a) to the report by Epstein (2012) of the relatively high pre-to-post test *gains* on the *Calculus Concept Inventory* [Epstein (2007)] by university students in a teacher centered course in Shanghai, China see “Re: FCI and CCI in China #2” [Hake (2012c)]. As a gauge of the attention garnered by Bao et al. (2009a), a Google <http://www.google.com/> search for [“Learning and Scientific Reasoning” Bao] (with the quotes but without the square brackets) turned up 16,800 hits on 31 January 2012 16:12-0800 at <http://bit.ly/wjjleL>. [*Caution*: the hit number has varied erratically from 6,150 to 47,500 over the period 14-31 January 2012. A Google Scholar <http://scholar.google.com/> search for [“Learning and Scientific Reasoning” Bao] on 31 January 2012 16:14-0800 netted 27 hits at <http://bit.ly/y5hMc0>.

Bao, L., K. Fang, T. Cai, J. Wang, L. Yang, L. Cui, J. Han, L. Ding, & Y. Luo. 2009b. “Learning of content knowledge and development of scientific reasoning ability: A cross culture comparison.” *Am. J. Phys*. **77**(12): 1118-1123; online to subscribers at <http://bit.ly/wV1wbe>.

Benford, R. & A.E. Lawson. 2001. “Relationships Between Effective Inquiry Use and the Development of Scientific Reasoning Skills in College Biology Labs” (Arizona State University, Tempe, AZ, 2001); Educational Resources Information Center (ERIC) accession no. ED456157; online as a 668 kB pdf at <http://1.usa.gov/wyLlmC>. A report on this and other work was later reported by Lawson et al. (2002) in the *Journal of College Science Teaching*.

Ben-Hur, M., ed. 1994. “On Feuerstein's Instrumental Enrichment: A Collection.” IRI/Skylight. According to ICELP's (2006) “Bibliography” this is a “Collection of papers written by an international group of researchers on theoretical and applied aspects of Instrumental Enrichment (IE). Evaluation of the empirical status of IE, comparison between IE and other cognitive education programs, application of IE with culturally different and minority students.”

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 233 kB pdf at <http://bit.ly/c7ZQlr>.

Bransford, J.D., A. L. Brown, & R.R. Cocking, eds. 2000. *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*, National Academy Press, online at <http://bit.ly/fVCQ6M>.

Brush. 2004. "Comments on the Epistemological Shoehorn Debate" *Science & Education* **13**: 197–200; the abstract and first page are online at <http://bit.ly/yhEGBF>.

Cakir, M. 2008 . "Constructivist approaches to learning in science and their implication for science pedagogy: A literature review," *International Journal of Environmental and Science Education* **3**(4): 193 - 206; online as a 225 kB pdf at <http://bit.ly/zTkk3S>.

Carlson, S.M. 2009. "Explaining the Reasoning-Fact Gap," Letters to the Editor, *Science* **323**: 1169; online at <http://bit.ly/zMscvL> - scroll to the bottom.

CAT. 2012. "Critical thinking Assessment Test" Tennessee Technological University, online at <http://www.tntech.edu/cat/home/>. According to the Overview:

> "The CAT Instrument is a unique tool designed to assess and promote the improvement of critical thinking and real-world problem solving skills. The instrument is the product of extensive development, testing, and refinement with a broad range of institutions, faculty, and students across the country. The National Science Foundation has provided support for many of these activities."

Clement, J. (Texas). 2009. "Learning and Scientific Reasoning - SCIENCE article Jan 30, 2009 pp. 586-7," PhysLrnR post of 5 Feb 2009 09:40:47-0600; online on the CLOSED! PhysLrnR archives <http://bit.ly/w5iA3p>.

Coletta & Phillips. 2005. "Interpreting FCI Scores: Normalized Gain, Preinstruction Scores, & Scientific Reasoning Ability," *Am. J. Phys.* **73**(12): 1172-1182; online to subscribers at <http://bit.ly/cRqwXS>. See also Coletta et al. (2007a,b) and Coletta & Phillips (2009).

Coletta, V.P., J.A. Phillips, & J.J. Steinert. 2007a. "Why You Should Measure Your Students' Reasoning Ability," *Phys. Teach.* **45**(4): 235-238; online to subscribers at <http://bit.ly/abH0YD>.

Coletta, V.P., J.A. Phillips, & J.J. Steinert. 2007b. "Interpreting force concept inventory scores: Normalized gain and SAT scores," *Phys. Rev. ST Phys. Educ. Res.* **3**, 010106, Issue 1 – June; online at <http://bit.ly/af0xQO>.

Coletta V.P., & J.A. Phillips. 2009. "Addressing Barriers to Conceptual Understanding in IE Physics Classes," *Physics Education Research Conference 2009*, Volume 1179: 117-120; online at <http://bit.ly/a7N6UO>.

COD. 2012. *Creative Organizational Design*, sells the "Watson-Glaser Critical Thinking Appraisal" online at <http://bit.ly/nM6lPe>.]

Darling-Hammond, L. & J. Bransford, eds. 2005. *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass. Publisher's information at <http://bit.ly/Azj2sk> .

Darling-Hammond, L. 2005. *A Good Teacher in Every Classroom: Preparing the Highly Qualified Teachers Our Children Deserve*." Jossey-Bass. Publisher's information at <http://bit.ly/A4aLMh>.

Darling-Hammond, L. 2010-2011. "Soaring Systems: High Flyers All Have Equitable Funding, Shared Curriculum, and Quality Teaching," *American Educator*, Winter, online as a 160 kB pdf at <http://bit.ly/A9EpUH>. Darling-Hammond wrote:

> "Although Finland, Singapore, and South Korea are very different from one another culturally and historically, all three have made startling improvements in their education systems over the last 30 years. Their investments have catapulted them from the bottom to the top of international rankings in student achievement and attainment, graduating more than 90 percent of their young people from high school and sending large majorities through college, far more than in the much wealthier United States. Their strategies also have much in common. All three:. . . . . *Organize teaching around national standards and a core curriculum that focus on higher-order thinking, inquiry, and problem solving through rigorous academic content*.. . . .[[My *italics*]]. . .. Working from lean national curriculum guides that have recommended assessment criteria, teachers collaborate to develop curriculum units and lessons at the school level, and develop school-based performance assessments—which include research projects, science investigations, and technology applications—to evaluate student learning."

Demetriou, A., J.-E. Gustafsson, A. Efklides, & M. Plastidou, 1994. "Structural systems in developing cognition, science, and education," pp. 81-106 in A. Demetriou, M. Shayer, & A. Efklides, eds., *Neo-Piagetian theories of cognitive development*. Routledge. Amazon.com information at <http://amzn.to/eb6R9S>, note the searchable "Look Inside" feature. An expurgated Google book preview is online at <http://bit.ly/fG9ws5>.

Dewey, J. 1916. "Method in Science Teaching," *General Science Quarterly* 1:3.

Ding, L., R. Chabay, B. Sherwood, and R. Beichner. 2006. "Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment," *Phys. Rev. ST Phys. Educ. Res.* **2**, 010105-1-7; online as a 301 kB pdf at <http://bit.ly/wrviCP>.

Einstein, A. 1936. *Physics and Reality*, J. of the Franklin Institute 221. An abstract is online at <http://bit.ly/hhfwwb>.

Elby, A., J. Fredriksen, C. Schwarz, and B. White. undated. "Epistemological beliefs assessment for physical science," online at <http://bit.ly/wow7jL>.

Elby, A. 1999. "Another reason that physics students learn by rote," *Am. J. Phys.* **67**(S1): S52-S57; online to subscribers at <http://bit.ly/zhzWd4>.

Epstein, J. 2007. "Development and Validation of the Calculus Concept Inventory," in *Proceedings of the Ninth International Conference on Mathematics Education in a Global Community*, 7-12 September, edited by Pugalee, Rogerson, & Schinck; online as a 48 kB pdf at <http://bit.ly/bqKSWJ>.

Epstein, J. 2012. "Re: FCI and CCI in China," PhysLrnR post of 13 Jan 2012 15:00:19-0500; online on the CLOSED! PhysLrnR archives at <http://bit.ly/xiLomm>.

Facione, P.A. 1990. “The California Critical Thinking Skills Test: College Level Technical Report #1 - Experimental Validation and Content Validity"; online as a 326 kB pdf at <http://1.usa.gov/qaWQFE> thanks to ERIC.

Feuerstein, R.. Y. Rand, M. Hoffman & R. Miller. 1980. *Instrumental Enrichment - An Intervention Program for Cognitive Modifiability*. ICLP Press. WARNING: Myron Tribus (2001f) wrote “AFTER taking a course of instruction from Feuerstein, I could read this book quite easily. Maybe it is my own limitation, but before taking his course I found the book difficult to read.” Tribus recommends starting with Ben-Hur (1994).

Fuller, R.G., T.C. Campbell, _D.I. Dykstra, & _S.M. Stevens. 2009. *College Teaching and the Development of Reasoning*. Information Age Publishing, publisher's information at <http://bit.ly/AddHRg>. Amazon.com information at <http://amzn.to/Ae6Qnp>.

Gerber, B.L., A.M. Cavallo, & E.A. Marek. 2001. “Relationships among informal learning, environments, teaching procedures, and scientific reasoning ability,” *Int. J. Sci. Educ.* **23**(5): 535-549; the abstract and first page are online at <http://bit.ly/yOow9O>.

Goffe, B. 2009. “FCI & BEMA Scores of Chinese Students in Bao, Cai, & Koenig,” PhysLrnR post of 14 May 2009 19:04:59-0400; online on the CLOSED! PhysLrnR archives at <http://bit.ly/zmnIGV>.

Gorder, P.F. 2009. “Study: Learning Science Facts Doesn’t Boost Science Reasoning,” Ohio State News Release, 29 January; online at <http://bit.ly/xuQAxd>.

Gross, P.R. 2009. “Learning Science: Content - With Reason,” *American Educator*, Fall, pp. 35-40; online as a 492 kB pdf at <http://bit.ly/fGnzMp> - note the stunning depictions of the beauty of nature, whose appreciation, like the comprehension of science, requires, according to Gross, *both* “content knowledge” and “reasoning.” See also Gross (2011).

Gross, P.R. 2011. “Review of the National Research Council's Framework for K-12 Science Education,” Thomas B. Fordham Institute <http://www.edexcellence.net/>; online as a 242 kB pdf at <http://bit.ly/oUfTZQ>.

Gutierrez, M. 2012. “Diane Ravitch shares her views on education reform,” *Sacramento Bee*, 24 January; online at <http://bit.ly/yNlElK>. Excerpts from Gutierrez’s interview of Ravitch (my *italics*):

On why she opposes using student test scores to evaluate teachers:

“I'm very dubious about these tests. First of all, because I know them. I spent seven years on the National Assessment Governing Board. I know the testing industry really well. I know how often the tests are flawed. And, I know what it does to kids when they are told they are a failure. Of what value is that?”

On her role as a vocal opponent of reform measures that she called "punitive" toward teachers:

“I think there has to be a far more thoughtful long-term plan to change the teaching profession and make it better than it is. It's not by demonizing teachers or firing teachers, but by first of all having much *more rigorous recruitment and higher standards for entry into teaching*. Secondly, by helping new teachers become better teachers and third, I think teacher evaluation is important, but not by test scores.”

On how to strengthen teacher training programs:

“*I think every teacher should have a strong liberal arts education* and should have very solid academic preparation for whatever they are going to teach or might teach. So, *if they are going to be a math teacher, they should have a bachelor's degree in mathematics*.”

On the role education schools can play in improving teacher quality:

“*They should probably be more selective* in terms of who gets in. That's not in their interest, because education schools have traditionally been the cash cow of the university. They take in everyone and spew out everybody.”

On why she is opposed to alternative paths to becoming a teacher:

“These alternative pathways produce people who either contribute to the revolving door (in the teaching profession) or are unprepared. Should there be alternative pathways into medicine? How about alternative pathways into becoming a pilot? *When you really care about something, you don't want alternative pathways*. You want pathways that are proven"”

Haack, S. 2003. *Defending Science—Within Reason: Between Scientism and Cynicism*, Prometheus Books, publisher's information at <http://bit.ly/deSlqh>. Amazon.com information at <http://amzn.to/avmBVz>, note the searchable "Look Inside" feature.

Hake R.R. & R. Wakeland. 1997. “ ‘What's F? What's m? What's a?’: A Non-Circular SDI-TST- Lab Treatment of Newton’s Second Law” in *Conference on the Introductory Physics Course*, Jack Wilson, ed. Wiley. pp. 277-283. [See also SDI Lab #6 “Newton's Second Law Revisited,” online at <http://www.physics.indiana.edu/~sdi>.]

Hake, R.R. 1987. “Promoting Student Crossover to the Newtonian World,” *Am. J. Phys.* **55**(10): 878-884; online as a 788 kB pdf at <http://bit.ly/a6vc3H>.

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>. See also Hake (1998b).

Hake, R.R. 1998b. “Interactive- engagement methods in introductory mechanics courses,” online at <http://bit.ly/aH2JQN>. A crucial companion paper to Hake (1998a). Submitted on 6/19/98 to the *Physics Education Research Supplement to AJP* (PERS), but rejected by its editor on the grounds that the very transparent, well-organized, and crystal-clear Physical-Review-type data tables were “impenetrable”!

Hake, R.R. 2000a. “The Need For Improved Physics Education of Teachers: FCI Pretest Scores for Graduates of High-School Physics Courses - Is it Finally Time To Implement Curriculum S?” Physics Education Research Conference 2000: Teacher Education, Univ. of Guelph, August 2-3; online as a 929 kB pdf at <http://bit.ly/gFvi8z>. For U.S. students enrolled in the Univ. of Indiana's non-calculus-based course for pre-meds and life-science majors in the Spring semesters of 1994 and 1995 it was reported that <pre> = 42% for 287 students who *had* take high-school physics and <pre> = 35% for 89 students who *had not* taken high-school physics.

Hake, R.R. 2000b. “The General Population’s Ignorance of Science Related Societal Issues: A Challenge for the University,” *AAPT Announcer* 30(2), 105 (2000); online as a 2.3 MB pdf at <http://bit.ly/9LxKOL>. Based on *The Science Illiteracy Crisis: A Challenge for The University,* a libretto for a Wagnerian musical drama to be presented under the stars in the open-air Santa Fe Opera Theater <http://www.santafeopera.org/>: an annotated interweaving of classic themes and original work, unpublished, 1989. The leitmotiv: “The road to U.S. science literacy begins with effective university science courses for pre-college teachers.”

Hake, R.R. 2001. “Socratic Dialog Inducing (SDI) Labs,” available online at <http://www.physics.indiana.edu/~sdi/>.

Hake, R.R. 2002a. “Lessons from the Physics Education Reform Effort,” *Ecology and Society* **2**: 28; online at <http://bit.ly/aL87VT>. See also “The Physics Education Reform Effort: A Possible Model for Higher Education”[Hake (2005)]. For an update on the six lessons on "interactive engagement" see Hake (2007).

Hake, R.R. 2002b. “Whence Do We Get the Teachers (Response to Madison),” PKAL Roundtable on the Future: *Assessment in the Service of Student Learning*, Duke University, March 1- 3; updated on 6/17/02; online as a 45 kB pdf at <http://bit.ly/w9M6dc>.

Hake, R. R. 2005. “The Physics Education Reform Effort: A Possible Model for Higher Education,” online as a 10 <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 2005, online to subscribers at <http://bit.ly/bvm8Ye> (if your institution doesn't subscribe to NTLF, then it should), and (b) disseminated by the *Tomorrow's Professor* list <http://bit.ly/9WAZ3Q> as Msg. 698 on 14 Feb 2006.

Hake, R.R. 2008. “Language Ambiguities in Education Research,” submitted to the insular *Journal of Learning Sciences* on 21 August but *rejected* - proof positive of an exemplary article; online at <http://bit.ly/bHTebD>. David Klahr <http://en.wikipedia.org/wiki/David_Klahr> wrote to me privately (quoted by permission):

> “I liked the paper. I think it's very thoughtful and nuanced. However it is tough going, even for someone as familiar with the issues (and as favorably cited by you) as I am. It's a shame that it was rejected, but I wonder if the reviewer just wasn't up to the very careful reading necessary to really follow your arguments all the way through. Even though I know this area quite well, obviously, I did have to really focus to fully understand the distinctions you were making.”

Hake, R.R. 2006a. “Possible Palliatives for the Paralyzing Pre/Post Paranoia that Plagues Some PEP's,” *Journal of MultiDisciplinary Evaluation*, Number 6, November, online as at 119 kB pdf at <http://bit.ly/caWtWl>. [PEP’s = Psychologists, Education specialists, Psychometricians]

Hake, R.R. 2006b. “Re: active learning needs a theory,” online on the OPEN! Phys-L archives at <http://bit.ly/d7QiET>. Post of of 2 Jun 2006 22:35:53-0700 to Phys-L and PhysLrnR. This is part of a 15-post Phys-L thread initiated by Josip Slisko (2006) on 2 Jun 2006 13:43:13-0500. See also Hake (2006c).

Hake, R.R. 2006c. “Re: active learning needs a theory,” online on the OPEN! AERA-L archives at <http://bit.ly/dB0mAU>. Post of 9 Jun 2006 13:00:45-0700 to AERA-L, PhysLrnR, and POD. This is part of the 16-post PhysLrnR thread initiated by Hake (2006b).

Hake, R.R. 2007. “Six Lessons From the Physics Education Reform Effort,” *Latin American Journal of Physics Education* **1**(1), September; online (with AIP style numbered references) as a 124 kB pdf at <http://bit.ly/bjvDOb>. Also available with APA style references as a 684 kB pdf at <http://bit.ly/96FWmE>.

Hake, R.R. 2009. “National Education Standards for the United States?” Online on the OPEN! AERA-L archives at <http://bit.ly/A8nllw>. Post of 9 Jun 2009 14:44:42-0700 to AERA-L

Hake, R.R. 2010a. "Education Research Employing Operational Definitions Can Enhance the Teaching Art," invited talk, Portland, OR, AAPT meeting, 19 July; online as a 3.8 MB pdf at <http://bit.ly/aGlkjm> and as ref. 60 at <http://www.physics.indiana.edu/~hake>.

Hake, R.R. 2010b. “Re: Metastudy on impact of inquiry in K-12,” AERA-L post of 3 Feb 2010 08:40:19 -0800; online on the OPEN! AERA-L archives at <http://bit.ly/zzrysV>.

Hake, R.R. 2011a. "Soak Time [was Academically Adrift?]" PhysLrnR post of 30 Jan 2011 16:49:56-0800; online on the CLOSED! PhysLrnR archives at <http://bit.ly/yAjjZH>.

Hake, R.R. 2011b. “The Impact of Concept Inventories On Physics Education and It's 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> and as ref. 64 at <http://bit.ly/a6M5y0>

Hake, R.R. 2011c. "Re: Cognitive Acceleration," AERA-L post of 14 Feb 2011 10:18:46 -0800; online on the OPEN! AERA-L archives at <http://bit.ly/xHKT91>. Gives 11 references to the work of Adey & Shayer on "Cognitive Acceleration," a program designed to promote student's thought processes from "concrete" to "formal" abstract thinking.

Hake, R.R. 2011d. "Feuerstein's Instrumental Enrichment," online on the OPEN! AERA-L archives at <http://bit.ly/htVa5a>. Post of 14 Feb 2011 16:33:00-0800 to AERA-L and Net-Gold. The abstract and link to the complete post were transmitted to various discussion lists are also online on my blog "Hake'sEdStuff" at <http://bit.ly/fASlF1> with a provision for comments.

Hake, R.R. 2011e. "The Overriding Influence of Poverty on Children's Educational Achievement" online on the OPEN! AERA-L archives at <http://bit.ly/tUU65W>. Post of 14 Dec 2011 09:56:02-0800 to AERA-L and Net-Gold. The abstract and link to the complete post were transmitted to several discussion lists.

.Hake, R.R. 2012a. "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 (2006a).

Hake, R.R. 2012b. "Helping Students to Think Like Scientists in Socratic Dialogue Inducing Labs," *Phys. Teach.* **50**(1): 48-52; online to subscribers at <http://bit.ly/wLy3En>. A version identical to the *Physics Teacher* article except for (a) minor formatting changes, and (b) the addition of a few hot-linked URL's is online at <http://bit.ly/x5ruYF>.

Hake, R.R. 2012c. "Re: FCI and CCI in China #2" online on the OPEN! AERA-L archives at <http://bit.ly/zz7WXk>. Post of 22 Jan 2012 16:27:43-0800 to AERA-L and Net-Gold. The abstract and link to the complete post are being transmitted to several discussion lists and are also on my blog "Hake'sEdStuff" at <http://bit.ly/AbW5oy> with a provision for comments.

Hake, R.R. 2012d. "Reform of STEM Education Requires A Systems Approach," online on the OPEN! AERA-L archives at <http://bit.ly/xcv0bR>. Post of 25 Jan 2012 11:12:56-0800 to AERA-L and Net-Gold. The abstract and link to the complete post are being transmitted to several discussion lists and are also on my blog "Hake'sEdStuff" at <http://bit.ly/A25rhg> with a provision for comments.

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 20 languages: Arabic, Chinese, Croatian, Czech, English, Finnish, French, French (Canadian), German, Greek, Italian, Japanese, Malaysian, Persian, Portuguese, Russian, Spanish, Slovak, Swedish, & Turkish.

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).

Hestenes D. & M. Wells. 1992. “A Mechanics Baseline Test,” Phys. Teach. 30: 159-166; online as a 283 kB pdf at <http://bit.ly/zWvRIK> [but without the test itself].

Hehn, J. & M. Neuschatz. 2006. “Physics for All? A Million and Counting!” *Phys. Today* **59**(2): 37; online to subscribers at <http://bit.ly/xX1Oif>.

Heiss, E.D., E.S. Osborne, & C.W. Hoffman. 1950. *Modern Science Teaching*. Macmillan.

IA. 2012. *Insight Assessment* at <http://www.insightassessment.com/>, sells the “Health Sciences Reasoning Test,” a test “developed for use by educators and researchers to assess the critical thinking skills of health science professionals and health science students.”

ICELP. 2006. International Center for the Enhancement of Learning Potential, online at <http://bit.ly/xQDJTZ>: “ICELP was established with the goal of continuing and expanding the educational and psychological work initiated by Prof. Feuerstein. The work of the ICELP is based on the theories of Structural Cognitive Modifiability and Mediated Learning Experience, which serve as a basis for three applied systems: the Learning Potential Assessment Device (LPAD), Instrumental Enrichment (IE) cognitive intervention program, and Shaping Modifying Environment. ICELP specializes in providing a wide range of Services, Trainings,
Research and Development.” See especially:

A. Entries under “Research” (top menu):
   (a) Feuerstein's and Other Theories <http://bit.ly/f3vMSa>;
   (b) "Aspects of Mediated Learning Experience" <http://bit.ly/giWcLi>;
   (c) Dynamic Cognitive Assessment <http://bit.ly/gcRVjL>;
   (d) Research on Instrumental Enrichment <http://bit.ly/f3vMSa>;
   (e) Research on Instrumental Enrichment <http://bit.ly/ghQgiQ>.

B. Bibliography under “Publications” (top menu):
   Bibliography of Books on the theories of MLE, LPAD and IE, their applied systems & research <http://bit.ly/yHLfEX>.

C. ICLP Press under “Publications” (top menu): <http://bit.ly/fVR1rH>. A good source of relatively up-to-date books on Feuerstein and his work.

JCSEE. 1994. Joint Committee on Standards for Educational Evaluation, “The Program Evaluation Standards,” 2nd ed., Sage. A glossary of evaluation terms from this publication online at <http://ec.wmich.edu/glossary/prog-glossary.htf>. This may eventually be replaced by terms from the 3rd edition of the Evaluation Standards by Yarbrough et al. (2011).

Kilpatrick, J., J. Swafford, B. Findell, eds. 2001. *Adding It Up: Helping Children Learn Mathematics*, National Academies Press, online at <http://bit.ly/xy4rGN>.

Klahr, D. & J. Li. 2005. “Cognitive Research and Elementary Science Instruction: From the Laboratory, to the Classroom, and Back,” *Journal of Science Education and Technology* **14**(2): 217-238; online as a 536 kB pdf at <http://bit.ly/apA7es>.

Labov, J.B., S.R. Singer, M.D. George, H.A. Schweingruber, & M.L. Hilton. 2009. "Effective Practices in Undergraduate STEM Education Part 1: Examining the Evidence," *CBE Life Sci Educ* **8**(3): 157-161; online at <http://bit.ly/cRc0JC>. This is a discussion of the "Workshop on Linking Evidence and Promising Practices in STEM Undergraduate Education" [National Academies (2008)].

Lasry, N., N. Finklestein, & E. Mazur. 2009. "Are Most People Too Dumb for Physics?" *Phys. Teach.* **47**(9): 418-422; online as a 233 kB pdf at <http://bit.ly/xK0oBJ>.

Lawson, A.E. 1978. "The development and validation of a classroom test of formal reasoning," *Journal of Research in Science Teaching* **15**(1): 11-24 (1978). See also Lawson (2009). In their *Am. J. Phys.* article [Bao et al. (2009b)] state in ref. 11 that they used Lawson's "Classroom Test of Scientific Reasoning" as revised in 2000, i.e., most likely the 24-item purely multiple-choice version of the test given in the Appendix of Coletta & Phillips (2005).

Lawson, A.E. 1982. "The Reality of General Cognitive Operations," *Sci. Ed.* **66**: 229 - 241; the first page is online at <http://bit.ly/bBYAN0>.

Lawson, A.E. 1995. *Science Teaching and the Development of Thinking*. Wadsworth. Amazon.com information at <http://amzn.to/aqWhQ2>. Appendix F contains the "Classroom Test of Scientific Reasoning," a 12-item test requiring written responses that, according to Lawson: (a) is a test of "ability to apply aspects of scientific and mathematical reasoning to analyze a situation, make a prediction, or solve a problem"; (b) total scores indicate the following levels of thinking in terms of questions correct: of 0-4 (0%-33%): empirical-inductive; 5-8 (42%-67%): transitional; 9-12 (75%-100%): hypothetical-deductive. A 24-item purely multiple-choice (MC) version of this test is in the Appendix of Coletta & Phillips (2005). As far as I know (please correct me if I'm wrong) Lawson has not indicated how total scores on the 24-question tests relate to reasoning level. However, one might assume that scaling applies such for the 24-question test total scores indicate the following levels of thinking in terms of percentage of questions correct 0-37%: empirical-inductive; 38%-71%: transitional; 72%-100%: hypothetical-deductive. The "generality of hypothetico-deductive reasoning" is maintained by Lawson (2000). For Stephen Brush's argument that hypothetical-deductive reasoning, by itself, cannot explain how new theories and discoveries are accepted in science, contrary to the view of Lawson (2003), see "Comments on the Epistemological Shoehorn Debate" [Brush (2004)].

Lawson, A.E., R. Benford, I. Bloom, M. Carlson, K. Falconer, D. Hestenes, E. Judson, M. Piburn, D. Sawada, J. Turley, & S. Wycoff. 2002. "Evaluating College Science and Mathematics Instruction: A Reform Effort that Improves Teaching Skills," *Journal of College Science Teaching* **XXX**(6): 388-393 online as a 365 kB pdf at <http://bit.ly/xXUWf0>.

Lawson, A.E. 2000. "The generality of hypothetico-deductive reasoning: Making scientific thinking explicit." American Biology Teacher **62**(7): 482-495 online as a 434 kB pdf at <http://bit.ly/cbERAG>.

Lawson, A.E. 2002, "What Does Galileo's Discovery of Jupiter's Moons Tell Us about the Process of Scientific Discovery?", *Science and Education* **11**: 1–24; an abstract is online at <http://bit.ly/yi00Yz>.Click on "Show Summary."

Lawson, A.E. 2003. “Allchin’s Shoehorn, or Why Science Is Hypothetico-Deductive,” *Science and Education* **12**: 331–337; the abstract and first page are online at <http://bit.ly/AftuTy>. The abstract reads:

> “Allchin’s critique of my analysis of Galileo’s discovery of Jupiter’s moons, and of my characterization of science as hypothetico-deductive, contains several factual and conceptual errors. Thus, contrary to his attempt to paint scientific discovery in terms of blind search and limited induction, a careful analysis of the way humans spontaneously process information. . . .[[see Lawson (2006)]]. . . . and reason supports a general hypothetico-deductive theory of human information processing, reasoning, and scientific discovery.”

Lawson, A.E. 2004. “A Reply to Allchin’s 'Pseudohistory and Pseudoscience',” *Science & Education* **13:** 599–605; online to subscribers at <http://bit.ly/cZqMVD>.

Lawson, A.E. 2006. “Points of View: On the Implications of Neuroscience Research for Science Teaching and Learning: Are There Any?,” *CBE – Life Science Education* 5(2): 111–117; online at online at <http://1.usa.gov/ygN6Cg>.

Lawson, A.E. 2010. Review of Fuller et al. (2009), *Sci & Educ* **20**: 99-102; online as a 106 kB pdf at <http://bit.ly/zwkOCz>.

Linn, M.C., H. Lee, R. Tinker, F. Husic, J.L. Chiu. 2006. “Teaching and Assessing Knowledge Integration in Science,” *Science* **331**: 1049–1050; online as a 197 kB pdf at <http://bit.ly/xPcabO>. Supporting online material is online as a 295 kB pdf at <http://bit.ly/zljUyP>.

Marek E.A. & A.M.L. Cavallo. 1997. *The Learning Cycle: Elementary School Science and Beyond*. Heinemann; publisher's information at <http://bit.ly/zH6aDv>. Amazon.com information at <http://amzn.to/AEAZZL>.

Marusic, M. & J. Slisko. 2012. “Influence of Three Different Methods of Teaching Physics on the Gain in Students' Development of Reasoning,” *International Journal of Science Education* **34**(2): 301-326; an abstract is online at <http://bit.ly/phQzSp>.

Mayer, R. 2004. “Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction,” *American Psychologist* **59**(1): 14–19; an abstract is online at <http://bit.ly/xmGYXT>.

McDermott, L.C., P.R.L. Heron, & P.S. Shaffer. 2005. “Physics by Inquiry: A research-based approach to preparing K-12 teachers of physics and physical science,” *Forum on Education of The American Physical Society*, Summer 2005 Newsletter; online at <http://bit.ly/yjoMRi>.

McDermott, L.C. and the Physics Education Group (University of Washington). 1996. *Physics by Inquiry*. John Wiley & Sons; author’s information at <http://bit.ly/zPFmkd>. Publisher’s information: volume 1 <http://bit.ly/xKIN36>, volume 2 <http://bit.ly/x8c9lS>. Amazon.com information: volume 1 <http://amzn.to/xUqwsa>, volume 2 <http://amzn.to/zfwjD6>.

McKinnon, J., & J. Renner. 1971. "Are colleges concerned with intellectual development?" *Am. J. Phys.* **39**(9): 1047-1052; online to subscribers at <http://bit.ly/9lLxPF>.

Minner, D.D., A. J. Levy, J. Century. 2009. "Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002"; online as a 201 kB pdf at <http://bit.ly/wdJq4R>. Minner et al. wrote:

> "It is difficult to exactly trace the first appearance of inquiry instruction, but it was born out of the longstanding dialogue about the nature of learning and teaching, in particular from the work of Jean Piaget, Lev Vygotsky, and David Ausubel. The work of these theorists was blended into the philosophy of learning known as constructivism (Cakir, 2008), which was then used to shape instructional materials. These kinds of constructivism-based materials are commonly classified under the moniker of inquiry-based and include hands-on activities as a way to motivate and engage students while concretizing science concepts. Constructivist approaches emphasize that knowledge is constructed by an individual through active thinking, defined as selective attention, organization of information, and integration with or replacement of existing knowledge; and that social interaction is necessary to create shared meaning, therefore, an individual needs to be actively engaged both behaviorally and mentally in the learning process for learning to take place (Cakir, 2008; Mayer, 2004). As constructivist approaches permeated much of educational practice in the 1970s, it became particularly prominent in science education through the focus on inquiry."

Moltz, D. 2009. "Blinding Them With Science," *Inside Higher Ed*, 30 January; online at <http://bit.ly/lUUf0o>. Caution: the link to the FCI is to the 1992 version and the link to the LCTSR doesn't lead directly to that test.

National Academies. 2008. *Evidence on Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education: Workshop on Linking Evidence and Promising Practices in STEM Undergraduate Education* online at <http://bit.ly/fAhNpA>: Meeting 1 of 30 June, online at <http://bit.ly/ciNwjQ>; Meeting 2 of 13-14 October containing commissioned papers online at <http://bit.ly/ceg1Bx>. See also the commentary on these workshops by Labov et al. (2009), and the NRC (2011a) book "Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics Education: Summary of Two Workshops."

NAGB. 2009. *Science Framework for the 2009 National Assessment of Educational Progress*, National Assessment Governing Board; online as a 3.7 MB pdf at <http://bit.ly/whsOAL>.

NCSE. 2008?. "Comparing TIMSS with NAEP and PISA in Mathematics and Science," online as a 291 kB pdf at <http://1.usa.gov/xwDsEF>

NRC. 1996. *National Science Education Standards*, National Academy Press; online at <http://bit.ly/jSuaQD>. See also NRC (2000).

NRC. 2000. "Inquiry and the National Science Education Standards: A Guide for Teaching and Learning," National Academy Press; online at <http://bit.ly/zM0tEP>. In the "Forward: A Scientists Perspective on Inquiry" to this report Bruce Alberts <http://en.wikipedia.org/wiki/Bruce_Alberts> defines "inquiry activities" as those that allow "students to conceptualize a question and then seek possible explanations that respond to that question."

NRC. 2011a. *Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics Education: Summary of Two Workshops* online at <http://bit.ly/nCMLk7>. Natalie Nielsen, Rapporteur. Planning Committee on Evidence on Selected Innovations in Undergraduate STEM Education: Susan Singer (Chair), Melvin George, Kenneth Heller, David Mogk, & William B. Wood, Board on Science Education, Division of Behavioral and Social Sciences and Education. The National Academies Press; online at <http://bit.ly/nCMLk7>. The description reads:

"Numerous teaching, learning, assessment, and institutional innovations in undergraduate science, technology, engineering, and mathematics (STEM) education have emerged in the past decade. Because virtually all of these innovations have been developed independently of one another, their goals and purposes vary widely. Some focus on making science accessible and meaningful to the vast majority of students who will not pursue STEM majors or careers; others aim to increase the diversity of students who enroll and succeed in STEM courses and programs; still other efforts focus on reforming the overall curriculum in specific disciplines. In addition to this variation in focus, these innovations have been implemented at scales that range from individual classrooms to entire departments or institutions.

By 2008, partly because of this wide variability, it was apparent that little was known about the feasibility of replicating individual innovations or about their potential for broader impact beyond the specific contexts in which they were created. The research base on innovations in undergraduate STEM education was expanding rapidly, but the process of synthesizing that knowledge base had not yet begun. If future investments were to be informed by the past, then the field clearly needed a retrospective look at the ways in which earlier innovations had influenced undergraduate STEM education.

To address this need, the National Research Council (NRC) convened two public workshops to examine the impact and effectiveness of selected STEM undergraduate education innovations. This volume summarizes the workshops, which addressed such topics as the link between learning goals and evidence; promising practices at the individual faculty and institutional levels; classroom-based promising practices; and professional development for graduate students, new faculty, and veteran faculty. The workshops concluded with a broader examination of the barriers and opportunities associated with systemic change."

NRC. 2011b. *Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, online at <http://bit.ly/zy0qqG>. For comments on this report see Paul Gross's (2011) "Review of the National Research Council's Framework for K-12 Science Education," and Helen Quinn's (2011) "A Framework for K-12 Science Education." Quinn is the chair of the NRC committed that produced the report. The NRC's description reads:

> "Science, engineering, and technology permeate nearly every facet of modern life and hold the key to meeting many of humanity's most pressing challenges, both present and future. To address the critical issues of U.S. competitiveness and to better prepare the workforce, *Framework for K-12 Science Education* proposes a new approach to K-12 science education that will capture students' interest and provide them with the necessary foundational knowledge in the field.
>
> *Framework for K-12 Science Education* outlines a broad set of expectations for students in science and engineering in grades K-12. These expectations will inform the development of new standards for K-12 science education and, subsequently, revisions to curriculum, instruction, assessment, and professional development for educators. This book identifies three dimensions that convey the disciplinary core ideas and practices around which science and engineering education in these grades should be built. These three dimensions are: cross-cutting concepts that unify the study of science and engineering through their common application across these fields; scientific and engineering practices; and core ideas in four disciplinary areas: physical sciences, life sciences, earth and space sciences, and engineering, technology, and the applications of science. The overarching goal is for all high school graduates to have sufficient knowledge of science and engineering to engage in public discussions on science-related issues; be careful consumers of scientific and technological information; and have the skills to enter the careers of their choice.
>
> *Framework for K-12 Science Education* is the first step in a process that will inform state-level decisions and provide a research-grounded basis for improving science teaching and learning across the country. The book will guide standards developers, curriculum designers, assessment developers, teacher educators, state and district science administrators, teachers, and educators who work in informal science environments."

Otero, V.K. & K.E. Gray. 2007. "Learning to Think Like Scientists with the PET Curriculum," in L. McCullough, J. Hsu, & P. Heron, eds., *Physics Education Research Conference Proceedings* (AIP, pp. 160-163); an abstract is online at <http://bit.ly/zWDdzX>.

Quinn, H. 2009. "What is science?" *Physics Today* **62**(7): 8-9, online as a 754 kB pdf at <http://bit.ly/y3P8SG>, scroll down.

Quinn, H. 2011. *APS News* **20**(10), Back Page: "A Framework for K-12 Science Education," online at <http://bit.ly/yuMnTl>. Helen Quinn is the chair of the NRC committed that produced the report.

Ravitch, D. 2011. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. Basic Books, publisher's information at <http://bit.ly/pN1NJo>. Amazon.com information at <http://amzn.to/pAjeZU>, note the searchable "Look Inside" feature.

Schmidt, W., G. Leroi, S. Billinge, L. Lederman, A. Champagne, R. Hake, P. Heron, L. McDermott, F. Myers, R. Otto, J. Pasachoff, C. Pennypacker, & P. Williams. 2011. "Towards Coherence in Science Instruction: A Framework for Science Literacy," online as a 856 kB pdf at <http://bit.ly/y1EURF>.

Schmidt, W.H., R. Houang, & S. Shakrani. 2009. *International Lessons About National Standards*, Thomas B. Fordham Insititute; online as a 1. 8MB pdf at <http://bit.ly/xPjmJ4 >.

Schoenfeld, A.H. 1988. “Learning To Think Mathematically: Problem Solving, Metacognition, and Sense-Making in Mathematics” in D. Grouws, ed. *Handbook for Research on Mathematics Teaching and Learning* (pp. 334-370); online as s 1.2 MB pdf at <http://bit.ly/AuCqAD>.

Science Daily. 2009. “College Freshmen In US And China: Chinese Students Know More Science Facts But Neither Group Especially Skilled In Reasoning,” 29 January; online at <http://bit.ly/wGvSmu>.

Shavelson, R.J. & L. Huang. 2003. “Responding Responsibly To the Frenzy to Assess Learning in Higher Education,” *Change Magazine*, January/February; online at <http://www.stanford.edu/dept/SUSE/SEAL/> // “Papers from Ongoing Research” // “Higher Education Assessment,” where “//” means “click on.” See also Shavelson (2009).

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 – click on “PDF” - 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>.

Shavelson, R.J. 2009. *Measuring College Learning Responsibly: Accountability in a New Era*. Stanford University Press, publisher’s information at <http://bit.ly/cfoLbM>. Amazon.com information at <http://amzn.to/i6zjbx>, note the searchable “Look Inside” feature.

Shayer, M. & P. Adey, eds. 2002. *Learning Intelligence: Cognitive Acceleration across the Curriculum from 5 to 15 Years*. Open University Press. Amazon.com information at <http://amzn.to/ApldsT>, note the searchable “Look Inside” feature.

Shulman, L. 1986. “Those who understand: knowledge growth in teaching.” *Educational Researcher* 15(2): 4-14; the first page is online at <http://bit.ly/xCOCtZ>.

Shulman, L. 1987. “Knowledge and teaching: foundations of the new reform,” *Harvard Educational Review* 57:1-22; an abstract is online at <http://bit.ly/idImgX>.

Slisko, J. 2006. "active learning needs a theory," online on the OPEN! Phys-L archives at <http://bit.ly/d7V2Eg>. Post of 2 Jun 2006 13:43:13-0500.

Tobias, S. & T.M. Duffy. 2009. *Constructivist Instruction: Success or Failure?* Routledge, publisher's information at <http://bit.ly/doVuHS>. Amazon.com information at <http://amzn.to/nybW9F>.

Tribus, M. 2001. “Will Our Educational System Be the Solution or the Problem?” online at <http://bit.ly/zhKl2f>. “A discussion of how to consider the entire educational system, preschool to adult education as a system - see “Reform of STEM Education Requires A Systems Approach” [Hake (2012d)] - setting of priorities regarding what to include and what to leave out, the relation between brain research and quality management practices in education.” Contains some good references to Feuerstein's work, including Feuerstein et al. (1980) and Ben-Hur (1994).

Willingham, D.T. 2007a. "Critical Thinking: Why Is It So Hard to Teach?" *American Educator* **31**(2): 8 -19, online as a 209 kB pdf at <http://bit.ly/zpk8wa>. See also Willingham (2007).

Willingham, D.T. 2007b. "How Knowledge Helps: It Speeds and Strengthens Reading Comprehension, Learning - and Thinking" *American Educator* **31**(2): 8 -19, online as as a 676 kB pdf at <http://bit.ly/zlvoxW>.

Willingham, D.T. 2007. *Cognition: The Thinking Animal*, 3rd ed. Pearson, publisher's information at <http://bit.ly/yHmcts>. Amazon.com information at <http://amzn.to/xSEMyR>.

Wivagg, D. & D. Allchin. 2002. "The Dogma of 'The' Scientific Method," *The American Biology Teacher* **64**(9): 645-646; online as a 152 kB pdf at <http://bit.ly/yEIGYT>.

Weinstock, R. 1961. "Laws of Classical Motion: What's F? What's m? What's a?" *Am. J. Phys.* **29**(10): online to subscribers at <http://bit.ly/yqMCqL>.

Wyckoff, S. 2001. "Changing the Culture of Undergraduate Science Teaching," *Journal of College Science Teaching*, 30(5): 306-312; online at <http://bit.ly/wRkZEp>. Fig. 2 shows results for 584 students in a "Bio 100" class that "demonstrates" a 41% gain in students reasoning abilities. The test used to gauge "reasoning abilities" is indicated to be "one of Lawson's reasoning tests." The "Bio 100" class data is among other data reported in Lawson et al. (2002) where it's implied that the reasoning test was the 24-question version of the *Lawson Classroom Test of Scientific Reasoning* [Lawson (1978)].

Yarbrough, D.B., L.M. Shulha, R.K. Hopson, & F.A. Caruthers. 2011. *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users*, Third Edition. Amazon.com information at <http://bit.ly/l98gg2>.

Zheng, A.Y., J.K. Lawhorn, T. Lumley, & S. Freeman. 2008. "Application of Bloom's Taxonomy Debunks the 'MCAT Myth': Analyses of questions that evaluate critical thinking, from college placement and medical school admission examinations, suggest improvements to college teaching methods." *Science* **319**: 414-415; online as a 201 kB pdf at <http://bit.ly/llpfgI>. For updates on "Bloom's Taxonomy" see Anderson & Sosniak (1994) and Anderson & Krathwohl (2001).

Zimmerman, C. 2007. "The development of scientific thinking skills in elementary and middle school," *Developmental Review* **27**: 172-223; online as a 729 kB pdf at <http://bit.ly/AlS6iw>. See also Zimmerman (2005).

Zimmerman, C. 2005. "The Development of Scientific Reasoning: What psychologists contribute to an Understanding of Elementary Science Learning," paper commissioned by the National Academies of Science (National Research Council's Board of Science Education, Consensus Study on Learning Science, Kindergarten through Eighth Grade); online as a 541 kB pdf at <http://bit.ly/cfWEoL>.