Education Research Employing Operational Definitions Can Enhance the Teaching Art*†◊

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

ABSTRACT: The operational definition of terms such as inquiry, discovery, and direct instruction is required if research findings are to be meaningfully conveyed to the education community and the general public. For example, Klahr & Nigam (2004) demonstrated the superiority of what they defined as direct instruction over what they defined as discovery learning. But their research was widely misinterpreted as showing that direct instruction in all its various forms was superior to discovery learning in all its various forms.

Then Kirschner, Sweller, & Clark [2006] added to the confusion by incorrectly: (a) identifying constructivist, discovery, problem-based, experiential, and inquiry-based teaching methods as all "minimally guided"; and (b) proclaiming them all to be failures.

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.”


†The reference is “Education Research Employing Operational Definitions Can Enhance the Teaching Art” [Hake (2010b)] – see below in References, pp. 18-26.

◊ See also “Helping Students to Think Like Physicists in Socratic Dialogue Inducing Labs” [Hake (2010a)] – see below in References, pp. 18-26.

© Richard R. Hake, 2 August 2010. Partially supported by NSF Grant DUE/MDR-9253965. Permission to copy or disseminate all or part of this material is granted provided that the copies are not made or distributed for commercial advantage, and the copyright and its date appear. To disseminate otherwise, to republish, or to place at another website (instead of linking to <http://www.physics.indiana.edu/~hake>) requires written permission. Please send comments and criticisms to <hake@earthlink.net>.
OUTLINE

I. What Are Operational Definitions?

II. Failure to Employ Operational Definitions For Pedagogical Terms Threatens the Teaching Art – Examples:


III. Epilogue

*References here and throughout this paper are below in References, pp. 18-26.
I. What Are Operational Definitions?

“All intelligent human endeavor stands with one foot on observation and the other on contemplation. But scientists have gradually come to limit themselves to certain types of observations and thought processes. One distinctly striking ‘limitation’ is the tacit desire among scientists to assure that they are in fact discussing the same concepts in a given argument. . . .

For example, the concept ‘length of an object’ as used in science is ultimately defined the very operation involved in making the measurement. The problem, ‘What is the length of a block,’ is for all practical purposes identical with the question, ‘What is the difference in between those two numbers printed on a specific meter stick that stand directly below two corresponding scratches, entered there to signify local coincidence with adjacent corners of the block?’ . . . .[[My insert: or, more simply - the reading on a vernier caliper whose arms bracket the length of the block (as demonstrated during my talk)]] . . . .

The bold italicized sentence above contains an abbreviated example of we shall call an operational definition, that of length, and although it seems ridiculously involved, and may never be used explicitly, four experts engaged to measure the length can regard it as the true meaning of the length of a block, available for examination should any dispute arise. Ideally, each of the concepts used in the physical sciences can be made clear in terms of some such operational definition, and that perhaps is the most important of the mechanisms whereby mutual understanding among scientists is made possible. For clearly it is more difficult to misinterpret actions than words. . . . ."

Similarly:

Henri Poincaré incisively commented:

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

Henri Poincaré (1905)

And in criticisms [Hake (2004, 2005) of the California Curriculum Commission’s anti-hands-on “Criteria For Evaluating K-8 Science Instructional Materials In Preparation for the 2006 Adoption,” I opined that popular pedagogic terms such as discovery learning, direct instruction, hands-on activities, active learning, cooperative learning, inquiry, and interactive engagement, should be operationally defined, i.e., rigorous operations should be defined for distinguishing pedagogic method X from other methods Y, Z, A, B, C, . . .

Consistent with:
“The true meaning of a term is found by observing what a man does with it, not what he says about it. . . . . I believe that many of the questions asked about social and philosophical subjects will be found meaningless when examined from the point of view of operations. It would doubtless conduce greatly to clarity of thought if the operational mode of thinking were adopted in all modes of inquiry.”

Percy Bridgman (1960)

Not far from the Logical Positivist credo:

“A statement is held to be literally meaningful only if it is either analytic or empirically verifiable.”

Alfred Jules Ayer (2001)

**WAIT! WAIT!! ANTI-POSITIVIST VIGILANTES - HOLD YOUR FIRE!!**
“Nowadays, the term *positivist* is widely used as a generalized term of abuse. As a literal designator it ceased to have any useful function – those philosophers to whom the term accurately applies have long since shuffled off this mortal coil, while any living social scientists who either bandy the term around or are the recipients of it as an abusive label are so confused about what it means that, while the word is full of sound and fury, it signifies nothing.
who realize nothing of this, still claim to see positivists everywhere. (When one is confused or suffering from delirium, it is possible to see anything.) Displaying what often amounts to an embarrassing degree of philosophical illiteracy, the vigilantes rarely bother to distinguish between classical (or Comtean) positivists . . . [[see, e.g., <http://en.wikipedia.org/wiki/Auguste_Comte>]]. . . ., on the one hand, and the more nefarious logical positivists, on the other. . . .[[see e.g., <http://en.wikipedia.org/wiki/Logical_positivism>]]. . . . Furthermore, they use a number of faulty criteria, either singly or in combination, to identify their illusory foe. The general fantasy is that anyone who is impressed by the sciences as a pinnacle of achievement of human knowledge, anyone who uses statistics or numerical data, anyone who believes that hypotheses need to be substantially warranted, anyone who is a realist (another unanalyzed but clearly derogatory word) is thereby a positivist.”

D.C. Phillips (2000, p. 157)
II. Failure to Employ Operational Definitions For Pedagogical Terms Threatens the Teaching Art – Examples:

A. "The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning" [Klahr & Nigam (2004)].

David Klahr and Milena Nigam (2004) demonstrated the superiority of what they defined as direct to what they defined as discovery learning. Although they were careful to operationally define their own very restricted meanings of direct instruction ["the goals, the materials, the examples, the explanations, and the pace of instruction were all teacher controlled"] and discovery learning ["no teacher intervention beyond the suggestion of a learning objective; no guiding questions; no feedback about the quality of the child’s selection of materials, explorations, or self-assessments"] their paper was widely presented in the media in ways that could be interpreted to imply that direct instruction in all its various forms was superior to discovery learning in all its various forms. For example:
“In science, how is critical thinking best taught? This question may be answered. . . . [by Klahr & Nigam, who]. . . . have new evidence that direct instruction - explicit teaching about how to design unconfounded experiments - most effectively helps elementary school students transfer their mastery of this important aspect of the scientific method from one experiment to another.”

Sean Cavanagh (2004) wrote in *Education Week*:
“The National Research Council is conducting a series of studies aimed at exploring topics such as the role of the laboratory in science classrooms and how states should assess students' knowledge in the subject. That renewed interest was also obvious with the release of a widely distributed study conducted by researchers at Carnegie Mellon University and the University of Pittsburgh, which was detailed at a national science “summit” sponsored by the U.S. Department of Education earlier this year. The study found that students taught through direct instruction were more likely on average to become “experts” in designing scientific experiments - an important step in the development of scientific-reasoning skills - than those taught through what is often called discovery learning. “ [My italics.]

Sharon Begley (2004a) wrote in the *Wall Street Journal*:
“It is conventional wisdom in science education. . . . that the best way to give K-12 students a deep and enduring understanding is through discovery learning . . . . the teacher gives the kids a goal and the requisite materials and then tells them to go to it, with the hope that they will uncover principles such as Newton's laws of motion. In contrast, using direct instruction, teachers explicitly present information to students. ‘The idea is that students who acquire knowledge on their own can apply it more broadly and extend it better than if they are told or shown that same knowledge,’ says David Klahr of Carnegie Mellon University in Pittsburgh. To test this claim, he and a colleague compared how well the approaches taught 112 third- and fourth-graders a core scientific concept: To discover how one thing affects another, change only one variable at a time. . . . . Students receiving direct instruction were explicitly told to change one property at a time and were given explanations. The discovery learners got neither. In both cases, the kids worked
with ramps and balls, so everyone did hands-on science. The result: Not only did more kids master the control-of-variables lesson from direct instruction, but -- and this strikes at the heart of the claims for discovery learning -- the latter approach did not give kids a deeper, more enduring knowledge.”

The American Association for the Advancement of Science in its *AAAS EurekAlert* (1998) stated:

“Carnegie Mellon Researchers Say Direct Instruction, Rather Than ‘Discovery Learning’ Is Best Way To Teach Process Skills In Science: Direct instruction using the Control of Variables Strategy, rather than discovery learning, may be the best way to teach young children about science, says a Carnegie Mellon psychologist who is conducting a four-year field study in public schools in Pittsburgh, Pa. The field study could lead to a new kind of science curriculum for elementary schools.”

It is ironic that the above sensational heading of the AAAS announcement, trumpeting the research of the *pro*-hands-on Klahr & Nigam (2004), merited a link on the virulently anti-reform *Mathematically Correct Science Center* website [MCSC (2008)], next to a link to the *anti*-hands-on testimony of California Curriculum Committee leader Stan Metzenberg (1998) before the U.S. House of Representatives.
In “Will the No Child Left Behind Act Promote Direct Instruction of Science?” [Hake (2005)]. I listed the widespread misinterpretation (see above) of Klahr & Nigam (2004), as one of the seven reasons why direct science instruction (in the passive-student sense – see below) might dominate K-12 science education under the aegis of the “No Child Left Behind Act.”

*Scholars at a Lecture* [William Hogarth (1822)]
Klahr & Li (2005), disturbed by the above reports, lamented [my insert at “...[[insert]]....”]; my **bold** text:

“Because of...[[media reports such as the above]]....others are concerned that our findings may be used to ‘conclude that direct instruction is the best way to teach science’ [Tweed (2004)], to promote lecture-based passive learning ‘Stand and deliver...or let them discover?’...[District Administration (2004)]...., and to equate our specification of discovery learning with the more moderate (and presumably, more often used) versions of guided or scaffolded inquiry....we share the concern that our findings may be misinterpreted as evidence to promote one method over another for science education as a whole...we are (now!) mindful of the way in which our results can be used to support or attack specific aspects of science education practice and policy....we may have muddied the interpretation of our findings by incorporating popular terminology like direct instruction and discovery learning into articles and public presentations of....[Klahr & Nigam (2004)].

Only when we tuned in to the recent political debate in California about the permissible amounts of hands-on science vs. direct instruction......[[Strauss (2004); Hake (2004, 2005); Woolf (2005)]].......did we become fully aware of how easy it is for someone to pick up a terminology, and imbue it with whatever meaning suits the purpose of an argument.......One thing is clear from all of this: **it is essential for the field of education to make much more precise use of terminology before moving on to public debates and policy decisions.** Indeed, it is surprising that when education researchers and science educators join in heated debates about discovery learning, direct instruction, inquiry, hands-on, or minds-on, they usually abandon one of the foundations of science - the operational definition. The field of science cannot advance without clear, unambiguous, operationally defined, and replicable procedures. Education science is no exception.”
Then Kirschner, Sweller, & Clark (2006) added fuel to their *direct-instruction-celebration* bonfire (while at the same time incinerating a *straw man* which they called *minimal guidance* with:


and fanned the fire by proclaiming Klahr & Nigam’s endorsement of *direct instruction* with this passage:

“Klahr and Nigam (2004) in a very important study, not only tested whether science learners learned more via a *discovery* versus *direct instruction* route but also, once learning had occurred, whether the quality of learning differed. Specifically, they tested whether those who had learned through discovery were better able to transfer their learning to new contexts. *The findings were unambiguous. Direct instruction involving considerable guidance, including examples, resulted in vastly more learning than discovery*. . . . [my *emphasis*]. . . . Those relatively few students who learned via discovery showed no signs of superior quality of learning.”
More generally Kirschner, Sweller, & Clark’s (KSC’s) abstract sets forth their thesis [my bold text]:

****************************************************
ABSTRACT: Evidence for the superiority of guided instruction is explained in the context of our knowledge of human cognitive architecture, expert–novice differences, and cognitive load. Although un-guided or minimally guided instructional approaches are very popular and intuitively appealing, the point is made that these approaches ignore both the structures that constitute human cognitive architecture and evidence from empirical studies over the past half-century that consistently indicate that minimally guided instruction is less effective and less efficient than instructional approaches that place a strong emphasis on guidance of the student learning process. The advantage of guidance begins to recede only when learners have sufficiently high prior knowledge to provide “internal” guidance. Recent developments in instructional research and instructional design models that support guidance during instruction are briefly described.
****************************************************

The above abstract, together with their provocative and seemingly non sequitur title “Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching,” suggests that KSC regard all constructivist, discovery, problem-based, experiential, and inquiry-based teaching as (a) “un-guided or minimally guided,” and (b) “failures.”

On the other hand, critics of KSC – see e.g., Part II “The Evidence for Constructivism” in Tobias & Duffy (2009), Hmelo-Silver et al. (2007), Schmidt et al. (2007) - point out that these methods are generally neither: (a) “un-guided or minimally guided,” nor (b) “failures.”

I agree with the KSC critics.

In my opinion, KSC are incinerating a Straw Man:
KSC’s Un-guided or Minimally Guided *Straw Man*
III. Epilogue

As an Epilogue I have lifted a few sentences from Thomas Duffy’s (2009) “Building Lines of Communication and a Research Agenda” in:

Therein Duffy wrote (my bold and italic text; my insert at “. . . [[insert]]. . . ”):

“Constructivism and the Design of Instruction: Success or Failure? That was the question underlying this volume. The not-so-surprising answer seems to be that success – or failure – is in the eye of the beholder. . . . . Unfortunately, Tobias and Duffy and most of the “beholders” who contributed to Tobias & Duffy (2009) appear to be either oblivious or dismissive of education research in STEM (Science, Technology, Engineering, & Mathematics) disciplines – see e.g., DeHaan (2005), Hake (1998a,b; 2002; 2008b), Heron & Meltzer (2005), Kelly et al. (2008), Klymkowsky et al. (2003), Michael (2006), Millis (2010), National Academies (2008), Nelson (2010), and Wood & Gentile (2003)). . . . the constructivist authors have argued consistently that guidance is essential; that it is simply a matter of the context of the guidance. And in an examination of the research described in the chapters and in the key journals, e.g., The Journal of the Learning Sciences. . . [in my opinion, more "key" for STEM disciplines are: e.g., Am. J. Phys. <http://bit.ly/d7xaDk>; J. Chem. Ed. <http://bit.ly/bwYYwJ>; J. Eng. Ed. <http://bit.ly/aW9epX>; Life Science Ed. <http://bit.ly/bmKAKL>; and Phys. Rev. Special Topics PER <http://bit.ly/9bCkUm>]. . . find numerous studies where there is considerable guidance, even direct instruction [Schwartz, Lindgren, & Lewis (2009)], and an emphasis on the need for a consistent structure (guidance) not only in the immediate instructional context but in the larger schooling context in which the instruction occurs [Herman & Gomez (2009), Duschl & Duncan (2009), Kolodner et al. (2003)].
INDEED THERE ARE NO STUDIES DESCRIBED IN THIS VOLUME WHERE GUIDANCE IS ABSENT.

The constructivists have their own myopic view, seemingly adverse to talking about mechanisms, in particular, information-processing mechanisms, that may underlie the effectiveness of guidance or scaffolding. As a consequence, the rationale for guidance and the prescription for guidance remains ill-defined. Klahr (2009) quite reasonably, calls for an operational definition of the various constructivist instructional models – definitions that would identify the key variables.

I can understand the frustration of Klahr and the misinterpretation of . . . [[Hey- Duffy! it’s by not of !!]]. . . . Kirschner (2009), Sweller (2009), and Mayer (2009) when the constructivist instructional approach is so ill-defined.

Indeed, perhaps the constructivists should design a learning environment to support learning about constructivism – certainly an ill-structured problem.”

A possible theme for the Physics Education Research Conference in 2011 (PERC 2011)?


Cavanagh, S. 2004. “‘NCLB Could Alter Science Teaching,’ Education Week 24(11): 1, 12-13, November 10; online at <http://tinyurl.com/2n4mrd>, scroll to the APPENDIX.

DeHaan, R.L. 2005. "The Impending Revolution in Undergraduate Science Education," *Journal of Science Education and Technology* **14**(2): 253-269; abstract online at <http://tinyurl.com/ymwwe3>. DeHaan wrote: “There is substantial evidence that scientific teaching in the sciences, i.e., teaching that employs instructional strategies that encourage undergraduates to become actively engaged in their own learning, can produce levels of understanding, retention and transfer of knowledge that are greater than those resulting from traditional lecture/lab classes. But widespread acceptance by university faculty of new pedagogies and curricular materials still lies in the future.”


Hake, R.R. 2008a. “Language Ambiguities in Education Research,” submitted to the *Journal of Learning Sciences* on 21 August but rejected – proof positive of an exemplary article; online at <http://bit.ly/bHTebD> and as ref. 54 at <http://www.physics.indiana.edu/~hake>. 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.”


Halloun, I. & Hestenes, D. 1985a. “The initial knowledge state of college physics,” Am. J. Phys. 53(11): 1043-1055; online at <http://modeling.asu.edu/R&E/Research.html>. ABSTRACT: “An instrument to assess the basic knowledge state of students taking a first course in physics has been designed and validated. Measurements with the instrument show that the student's initial qualitative, common sense beliefs about motion and causes has a large effect on performance in physics, but conventional instruction induces only a small change in those beliefs.” Contains the “Mechanics Diagnostic” test (omitted from the online version), precursor to the widely used “Force Concept Inventory” [Hestenes et al. (1992), Halloun et al.(1995)].


Holton, G. & S.G. Brush. 2001. *Physics the Human Adventure: From Copernicus to Einstein and Beyond*. Rutgers University Press, publisher’s information at <http://bit.ly/bqG4p7>. Amazon.com information at <http://amzn.to/9air1l>, note the searchable “Look Inside” feature. A Google “book preview” is online at <http://tinyurl.com/2nfts6>. Operational definitions are discussed in Chapter 12 “On the Nature of Concepts.” Holton and Brush add this clarification to their discussion of *operational definitions*: “It is necessary to enter here a warning against carrying this mode of thinking too far. Rigorous application of operationalism can be and in fact has been used to attack speculations that may well develop and turn out to be fruitful after all. For example, at the end of the nineteenth century the physicist Ernst Mach argued that the *atom* is a physically meaningless concept, because there was at that time no way to observe or measure its discrete individual properties. Nevertheless, atomic theories being developed in those years . . . . were to be of great value to science even though the atom at that time had to be defined in terms of observable physical and chemical interactions rather than in terms of measurable dimensions. Another attempt to distinguish between scientific and non-scientific statements was made by Karl Popper. . . . [[<http://en.wikipedia.org/wiki/Karl_Popper>]] . . . in his influential *The Logic of Scientific Discovery* [Popper (1934)]. Basing himself on his principle that science proceeds by making bold conjectures and attempting to refute them (Section 3.1) he declared: a statement that cannot be ‘falsified’ (that is, proven to be false) by any conceivable experiment is not a scientific statement. . . . Popper’s criterion may be too strict, if we interpret it to mean (as Popper himself originally did) that a theory is scientific only if it leads to testable predictions about events or phenomena not yet observed. The domain of science cannot be – and in practice has not been - limited to what can be brought into a terrestrial laboratory for controlled experimentation. The falsifiability criterion has been used, quite absurdly to argue that Darwinian evolution is not a scientific theory because we cannot use it to predict what species will evolve in the future. . . . *The essential point that both Bridgman and Popper wanted to make, and which must not be obscured by any misapplications of their criteria, is that scientific concepts must eventually be connected with observations about the real world if they are to survive. . . .[[My emphasis]] . . .” [In my opinion, the operationally undefined and unfalsifiable “isms” in much of the education literature allow little connection with the real world of classroom teaching.


Metzenberg, S. 1998. Testimony before the U.S. House of Representatives; online at **Mathematically Correct** <http://www.mathematicallycorrect.com/>, scroll down to and click on "Science Corner" under "Site Index" and then click on (a) "Stan Metzenberg at the House Science Committee" to bring up <http://mathematicallycorrect.com/stanmetz.htm>, and (b) “Follow-Up Questions for Dr. Stan Metzenberg” to bring up <http://mathematicallycorrect.com/moremetz.htm>.

Michael, J. 2006. “Where's the evidence that active learning works?” *Advances in Physiology Education* **30**: 159-167, online at <http://tinyurl.com/ykzp7lt>. Michael wrote: “One of the most striking findings [came from comparison of the learning outcomes (as measured by the FCI and a related inventory on mechanics) from 14 traditional courses (2,084 students) and 48 courses using "interactive-engagement" (active learning) techniques (4,458 students)] . . . . [Hake (1998a,b)] . . . . The results on the FCI assessment showed that students in the interactive engagement courses outperformed students in the traditional courses by 2 SDs. Similarly, students in the interactive-engagement courses outperformed students in the traditional courses on the Mechanics Baseline Test, a measure of problem-solving ability. This certainly looks like evidence that active learning works! Research in physics education is having a profound effect on the development of instructional materials.”


Tweed, A. 2004. “Direct Instruction: Is It the Most Effective Science Teaching Strategy” NSTA Reports, 15 December; response to Cavanagh (2004); online at <http://tinyurl.com/3a63x5>, scroll to the APPENDIX. At the time, Tweed was the president of the National Science Teachers Association (NSTA).

Wood, W.B., & J.M. Gentile. 2003. "Teaching in a research context," Science 302: 1510; 28 November; a summary is online at <http://bit.ly/9qGR6m>. Wood & Gentile wrote: “Physics educators have led the way in developing and using objective tests to compare student learning gains in different types of courses, and chemists, biologists, and others are now developing similar instruments. These tests provide convincing evidence that students assimilate new knowledge more effectively in courses including active, inquiry-based, and collaborative learning, assisted by information technology, than in traditional courses.”