Learning Goals in Undergraduate STEM Education and Evidence for Achieving Them

Jose P. Mestre
Departments of Physics and Educational Psychology
University of Illinois at Urbana/Champaign

In this white paper I propose a short list of learning goals in undergraduate STEM education and suggest the types of evidence that would indicate whether or not the learning goals are being achieved. Both the learning goals and proposed evidence will be accompanied by arguments and discussions about the relevance of the proposed goals given today's context. I also discuss why certain types of evidence should carry more weight than others, where current gaps in evidence exist, and why the quality of evidence is pivotal in promoting the adoption of promising instructional practices in undergraduate STEM instruction.

Setting the Context

The learning goals and evidence for achieving them that I will propose are motivated by the overarching goal of preparing a STEM workforce that is nimble enough to deal with the challenges of the present and future. I begin by making three observations to set the context.

Observation 1: Knowledge and information in the STEM fields is growing at an increasingly brisk pace. This knowledge/information explosion makes it impossible to "cover" in an undergraduate STEM course all of the important ideas/concepts/discoveries/facts.

Observation 2: An ever increasing number of interesting problems in the STEM fields now lie at the interface of multiple disciplines. Whereas in the past, it was possible for STEM professional to stay totally within discipline "silos" during their careers, tomorrow's workforce will need to be knowledgeable across multiple disciplines to grapple with the complex scientific and technological problems that society will need to solve.

Observation 3: Students in undergraduate STEM courses are being asked to "learn" an increasing body of knowledge only to forget it shortly after the courses are over. Instructional methods that help students retain and apply major concepts within and across STEM disciplines remain a major challenge in STEM education.

Proposed Learning Goals in STEM Undergraduate Education

Guided by the three observations above, I propose three learning goals, the first is about content coverage; the second is about promoting learning that lasts; the third is about helping students organize their knowledge for optimal efficiency in future learning and application.

Goal 1: *Structure instruction to help students learn a few major principles/concepts well and in-depth*. Given Observation 1, we will never be able to cover all, or even enough, of the important ideas in any undergraduate STEM course. Therefore, rather than worry about cramming more material into an already bloated curriculum, it would be best to focus on teaching a few of the major concepts/principles well in order to help students see "the big picture." Here teaching "well and in-depth" means that at the end of a course students should be able to verbalize the major ideas learned and discuss intelligently how those ideas relate to other related ideas and to the multiple contexts in which they can be applied (both within and across STEM disciplines), as well as apply those major ideas to analyze/solve problems.

Goal 2: *Structure instruction to help students retain what they learn over the long term*. Learning a few things well and in-depth should also mean learning them in a way that is not forgotten shortly after a course is over. Newly learned knowledge is quickly forgotten without the conceptual underpinnings on which to hang the newly learned knowledge.

Goal 3: *Assist students in building a mental framework that serves as a foundation for future learning*. Research into expert-novice differences has shown that possessing a mental framework, organized hierarchically by the big ideas and containing both conceptual and procedural knowledge, will facilitate organizing, remembering and recalling of new concepts (for overviews, see: Bransford, Brown, & Cocking, 1999; Etkina, Mestre, & O'Donnell, 2005). Thus, instruction should be designed to help students transfer knowledge flexibly, and to help students retain knowledge over the long term. Because of the increasing interdisciplinarity that will be required to tackle the interesting problems that will face us in the future today's students will need to become flexible learners capable of being retrained for the many jobs they will hold during their lifetimes. Preparing students with a knowledge base that is fundamental, that can be applied flexibly across contexts, and that is retained in the long term, will serve the student well in tomorrow's STEM world.

Note that I am not proposing a long list of goals—only three in fact—and that they are intended to help us think broadly about both what is important in STEM education, and how we should structure teaching and learning across the STEM disciplines to train the workforce of the future.

Evidence Needed to Conclude Learning Goals 1-3 Are Being Achieved

Having identified three broad learning goals, the question now becomes how we determine whether they are being reached. That is, what type of evidence is needed to conclude that progress is being made? What I will propose in this section moves somewhat into uncharted territories: Some of the types of assessments needed to produce the necessary evidence are non-traditional, or put a different way, current assessment practices are not geared to produce some of the type of evidence that I will argue is needed. I now present the type of evidence that I believe is necessary to judge whether or not the above three goals are being achieved.

Evidence Needed for Goal 1. Goal 1 has two components—depth and breadth. Students should be able to know a few of the major concepts/principles deeply, and they should be able to apply those flexibly across more than one STEM discipline. Thus, two types of evidence corresponding to these two components are proposed: 1) *Ability to display understanding of the conceptual underpinnings behind problem solutions*, and 2) *Ability to apply big ideas in relevant contexts both within a domain and across domains*.

<u>Commentary on item 1</u>: In quantitative STEM fields like physics, I find that students are able to perform amazingly well at the procedural level (i.e., manipulating equations to get answers) and at factual recall (at least in the short term), but they are dismal at being able to discuss the concepts/principles underlying problem solutions. This situation is likely fueled by current assessment practices since what is tested in typical homework and exams is largely procedural knowledge, which inadvertently discourages a focus on conceptual development if students' goal is to maximize their course grade. Unless conceptual and procedural knowledge are integrated, however, retention will remain illusive since there is no mental framework on which students can hang procedures and facts without the underlying meaning (concepts/principles). Assessment practices do not normally deal with testing students' conceptual understanding in depth, largely because it is much easier (and cheaper if the goal is to test large portions of a population) to assess factual and

procedural knowledge, especially with multiple choice tests. This only means that it will be difficult to develop assessment prototypes to measure depth of conceptual understanding in STEM undergraduate courses, not that it is impossible to do so.

I provide one example from my own work for adding more “depth” to introductory physics instruction for illustrative purposes. Over a decade ago I decided to teach a large introductory physics course in a way that emphasized the conceptual meaning behind problem solving by requiring what I called *strategy writing*. (Leonard, Dufresne & Mestre, 1996). The idea behind strategy writing was based on the findings of a classic study by Chi, Feltovich and Glaser (1981); they found that when experts were asked for the approach they would use to solve a problem they mentioned the major principle/concept, the justification for why the principle/concept applied to the particular context, and a general procedure for applying the principle/concept. In contrast, novices mentioned the equations they would use. In the course I taught, students were required to write “strategies” prior to generating solutions in all homework problems they turned in for grading and in some problems in hour exams. Strategies, students were told, were prose paragraphs (no equations allowed) that discussed in a coherent way the major principle/concept to be applied, the justification for why it applied to the problem, and a procedure for applying the principle/concept. When students asked for more details regarding how to write a strategy, they were simply told that there were many ways to write good strategies so long as they contained the three components, and that the best litmus test of the quality of a strategy was that if a good strategy was given to a student who was stuck in solving a particular problem, that student should be able to use it to generate a solution. Following the course, students who practiced strategy writing were significantly better at categorizing problems according to the major principle needed for solution compared to students enrolled in a traditional course, which is a trait of expertise (Chi, Feltovich & Glaser, 1981; Hardiman, Dufresne & Mestre, 1989). Strategy writing students were also much better than their non-strategy writing counterparts at retaining the big ideas learned in the course months after it was over.

<u>Commentary on item 2</u>: Whereas item 1 focuses on depth of understanding based on principled knowledge, item 2 focuses on ability to transfer knowledge flexibly (breadth of knowledge). Transfer of learning has been difficult to achieve in education, with much effort and debate going towards figuring out the conditions under which transfer is facilitated (see Ch 3 of Bransford et al., 1999; Detterman & Sternberg, 1993; Lobato, 2003, 2006; Marton, 2006; Mestre, 2003, 2005b). Research suggests that transfer is facilitated by teaching in ways that illustrate the multiple contexts in which major ideas apply and that encourage students to be active participants in their learning by verbalizing and reflecting on why and how major ideas are used to analyze and solve problems.

Evidence Needed for Goal 2. Goal 2 targets retention, or what I like to call "learning that lasts." I find it interesting that little effort goes into evaluating which instructional practices are better for helping students retain the knowledge they learn in STEM courses over the long term. It would seem that maximizing retention is fundamental to our educational mission, yet we know next-to-nothing about what students retain from STEM courses weeks, months, or years after completing courses. In psychology, some experiments have shown that flexible application of knowledge is facilitated when the knowledge is learned conceptually (rather than rotely). (References coming).

Evidence Needed for Goal 3. Goal 3 targets the formation of a highly organized mental network that facilitates knowledge application as well as new learning. Evidence for this goal is, I believe, the most difficult to acquire. Any evidence for how well structured in memory one's domain knowledge is will have to be deduced indirectly. There are some studies that provide some guidance. For example, Anderson and Shifrin (1980) demonstrated that children who had considerable expertise about spiders were better able than non-expert children at understanding and recalling the salient features of a passage about spiders. Another study by Voss, Vesonder, & Spilich (1980) demonstrated that experts in the game of baseball (fans, not players) were able to recall the important features from a passage about a baseball game, whereas non-experts recalled nonessential details about the game. Recent work in visual cognition also shows that experts in the game of American football are better able to detect meaningful changes made to football game scenes compared to non-experts (Werner & Thies, 2000). One of my recent studies also shows that physics experts are able to detect changes made (secretly) to diagrams while they are explaining the physics underlying the diagrams only if the changes modify the physics of the situation, whereas novices are unable to detect both physics-modifying and physics-non-modifying changes (Feil & Mestre, in review). These studies all suggest that those who know more and who have more nuanced understanding of a domain process situations in that domain in terms of the conceptual underpinnings, which in turn means that they are able to learn more efficiently from the situation. Thus, ability to learn new information efficiently, and ability to cue on meaningful features in a situation can be used as proxies for ascertaining the degree to which students are forming a hierarchically organized mental framework of the domain that is conducive for future learning.

Types of Evidence that Are More Compelling and Should Carry More Weight

The three goals and proposed evidence outlined above have a “motherhood and apple pie” quality about them. If asked, I do not think scientists would reject the notion that students in their classes should learn major concepts well and in depth, or conversely accept the notion that students in their classes should only be able to recall facts and learn procedures for solving problems. Nevertheless, the goals and evidence proposed do not drive instruction or educational policy in undergraduate STEM education at the present time. If there is agreement in the three observations made at the beginning of this white paper, and in the need for a STEM work force made up of nimble, lifelong learners of what will increasingly become interdisciplinary knowledge, then the need to teach in ways that promote principled knowledge is pivotal. Note that I am not arguing that factual and procedural knowledge in STEM domains is unimportant—it is important. I am arguing that if this is all we test for, then it will be all students that will pay attention to, and further that without a conceptual framework into which it can be integrated, factual and procedural knowledge will be quickly forgotten.

Existing Gaps in Evidence

Because assessment practices are heavy on factual recall and reproducing standard problem solving procedures, and light on measures of conceptual understanding, there needs to be a shift in emphasis in assessment practices to provide the evidence needed to judge whether progress is being made. Further, studies of transfer and of retention are rare in STEM, and hard to do well. There is a "chicken or the egg" dilemma at work here. To devise ways to provide the needed evidence requires knowledge of both innovative assessment strategies and cognition, but STEM professionals in the academy are not trained in these areas (for a more thorough discussion of this dilemma, see Etkina, et al., 2005).

In short, a major contributor to the gap in needed evidence is the lack of training in STEM Ph.D. programs in cognition, experimental design in education, and assessment. I am not advocating that STEM graduate students take a double course load in cognition/education and in their STEM discipline in route to their doctorates, but adding two to three courses to help them think about teaching based on cognitive models of the learner in order to promote retention and transfer, and about evaluating their teaching in terms of learning goals would provide the needed infrastructure to implement the goals proposed here.

What is certainly clear is that there appears to be no natural mechanism in undergraduate STEM programs to evaluate long-term retention of knowledge. Professors teach their courses, construct tests for what they deem important, and assign grades based on performance on those tests, but they remain blissful about what their students still know a few months after the course is over. Evidence of what students do or do not retain, and/or transfer across STEM undergraduate courses remains anecdotal; physicists complain that students in their introductory courses know little of the calculus they took in math classes; engineers complain that students do not see how the physics they learned is relevant in more applied engineering contexts. One question that needs to be addressed is: Whose responsibility is it to determine how much students retain months after taking STEM courses, and how will this information be fed back to improve STEM instruction?

Importance of Quality of Evidence for Adoption of Promising Instructional Practices

Shifting directions in undergraduate STEM education is akin to changing course for a large ocean vessel—it happens slowly. The evidence needed to effect change needs to be compelling in the sense of being valued by the profession as important, it needs to be of high quality, and it needs to be abundant. I will recount here a short version of how educational innovation was catalyzed in physics (for a more thorough discussion, see Mestre, 2005a).

Three decades ago physics university professors believed that if they presented material clearly enough, students would learn it. Evidence began to emerge from physics education research that students left courses taught by the most competent and charismatic of instructors with major misconceptions about the basic behavior of the physical world—misconceptions that were deeply seated and difficult to dislodge (see Clement, 1982; McDermott, 1984; Mestre, 1991, and references therein). The most compelling evidence for the presence of misconceptions after students finished physics courses came about largely by the development of a test called the Force Concept Inventory (FCI) developed by David Hestenes and his collaborators, a test that was based on research findings on students' misconceptions (Hestenes, Wells & Swackhammer, 1992). Shortly after the FCI was developed professors of introductory physics were hesitant to administer it to their classes. They reasoned that it would be a waste of class time to administer the FCI since it contained very simple ideas that they had taught well and surely their students knew this material. Further coaxing led

some professors to administer the FCI to their classes, and to their surprise, their students also showed poor understanding of basic concepts. Word spread, and more professors administered the FCI to their students with similar results. In turn, this newly found evidence led to curricula designed to help students overcome stubborn misconceptions (Camp & Clement, 1994; Laws, 1991; McDemortt,& Shaffer, 2003; Sokoloff & Thornton, 1997) and to meta-analyses of the types of instruction that resulted in better FCI performance (Hake, 1998); the Hake analysis indicated that pedagogies of engagement were much better at helping students overcome stubborn misconceptions than passive lectures.
It is noteworthy that the process of developing assessments to collect quality evidence, of using evidence to design instructional interventions that then had to be evaluated for effectiveness, took three decades.

The good news is that many other STEM disciplines are developing their version of the FCI (see: http://www.foundationcoalition.org/home/keycomponents/concept/index.html). The bad news is that this is a very narrow focus and we should look beyond identifying and eradicating misconceptions as the goal of STEM instruction, but at least it starts a much needed dialog toward curricular improvement.

Final Thoughts

The topic of learning goals and evidence for achieving them can be addressed at various grain sizes. I have chosen a course grain size, offering a small number of broad goals and guidance for the type of evidence needed to judge progress.

It is my hope that this will lead to thinking and discussions in the higher education STEM community about what should be the broad, fundamental goals in STEM education for the 21^{st} century. This conscientious attempt to put the horse before the cart is motivated by the belief that without first reaching some consensus about the fundamental goals of STEM higher education, we might be tempted to move directly to detailed goals at finer grain sizes, which in turn would lead to incremental changes; although finer grain sizes will be needed in this effort eventually, first we need to answer the question: What should undergraduate STEM education be about?

References

Anderson, R.C. & Shifrin, Z. (1980). The meaning of words in context. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), Theoretical Issues in Reading Comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bransford, J.D., Brown, A.L. & Cocking, R.R. (1999). How People Learn: Brain, Mind, Experience, and School. Washington, D.C.: National Academy Press.

Camp, C. & Clement J. (1994). Preconceptions in mechanics: Lessons dealing with students' conceptual difficulties. Dubuque, Iowa: Kendall/Hunt.

Clement, J.J. (1982). Students' preconceptions in introductory mechanics. American Journal of Physics, 50, 66-71.

Chi, M.T.H., Feltovich, P.J. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Detterman, D.K. & Sternberg, R.J. (Eds.) (1993). *Transfer on trial: Intelligence, cognition, and instruction*. Norwood, NJ: Ablex.

Etkina, E., Mestre, J., & O'Donnell, A. (2005). The impact of the cognitive revolution on science learning and teaching. In J.M Royer (Ed.) *The Cognitive Revolution in Educational Psychology* (pp. 119-164). Greenwich, CT: Information Age Publishing.

Feil, A. & Mestre, J. (in review). Change blindness among experts and novices in physics.

Hake, R. (1998). Interactive-engagement vs traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, American Journal of Physics, 66, 64-74.

Hardiman, P.T., Dufresne, R. & Mestre, J.P. (1989). The relation between problem categorization and problem solving among experts and novices. Memory & Cognition, 17, 627-638.

Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force concept inventory. The Physics Teacher, 30, (March), 159-166.

Laws. P. (1991). Calculus-based physics without lectures. Physics Today, 44, #12 (Dec.), 24-31.

Leonard, W.J., Dufresne, R.J., & Mestre, J.P. (1996). Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in solving problems. American Journal of Physics, 64, 1495-1503.

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher*, *32*(1), 17-20.

Lobato, J. (2006). Alternative perspectives on the transfer of learning: History, issues, and challenges for future research. Journal of the Learning Sciences, 15(#4), 431-449.

Marton, F. (2006). Sameness and difference in transfer. Journal of the Learning Sciences, 15(#4), 501-535.

McDermott, L.C., Shaffer, P.S.,& the Physics Education Group at the University of Washington (2003). Tutorials in Introductory Physics. Upper Saddle River, NJ: Prentice Hall.

McDermott, L.C. (1984). Research on conceptual understanding in mechanics. Physics Today, 37(7), 24-32.

Mestre, J. (2003). *Transfer of Learning: Issues and Research Agenda*. National Science Foundation Report #NSF03-212. 27 pages. (Available at www.nsf.gov/pubs/2003/nsf03212/nsf03212.pdf.)

Mestre, J.P. (2005a). Facts and myths about pedagogies of engagement in science learning. Peer Review, 7(#2, Winter Issue), 24-27.

Mestre (Ed.) (2005b), *Transfer of Learning from a Modern Multidisciplinary Perspective*. Greenwich, CT: Information Age Publishing.

Sokoloff, D.R., & Thornton, R.K. (1997). Using interactive lecture demonstrations to create an active learning environment. The Physics Teacher, 35 (6), 340-347.

Mestre, J.P. (1991). Learning and instruction in pre-college physical science. Physics Today, 44, #9 (Sept.), 56-62.

Werner, S., & Thies, B. (2000). Is “Change Blindness” attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*, *7*, 163-173.

Voss, J.F., Vesonder, G.T. & Spilich, G.J. (1980) Text generation and recall by high knowledge and low knowledge individuals. Journal of Verbal Learning and Verbal Behavior, 19, 651-667.