A new model for post-secondary education, the Optimized University

Carl Wieman
Dept. of Physics
University of Colorado and University of British Columbia

I. Introductory caveats.

This article is discussing the situation at North American universities in general, and not specifically those in BC. Some details may be slightly different in BC from the rest of the continent, as well as varying between different institutions of higher education in BC. However indications are that most of the general features are largely the same throughout North America. For example, the National Survey of Student Engagement is an extensive survey of students' college educational experience that covers institutions of higher education in both the US and Canada; the responses are quite similar in the two countries. Also, there is considerable movement of faculty and administrators between and throughout the two countries.

My remarks will also be limited to discussing higher education in the natural and applied sciences at the undergraduate level. This is an important but intentionally quite limited aspect of higher education. Far too often, discussions of higher education attempt to deal with the entire vast subject, and as a result the conclusions are so general they can offer little useful guidance. I am limiting my remarks according to both discipline (pure and applied sciences), function (undergraduate education), and institutional type (research university) in the hope this can provide useful guidance for specific actions and changes at the level where actions can actually take place.

Although I will focus on undergraduate science education, I do not mean to imply that there are not many other important roles and functions of higher education. In addition to other educational goals, there also are other major functions fulfilled by the modern university, most notably research and the closely linked graduate education, as well service to the communities and nation via support of the economy, hospitals connected with medical schools, etc. In this article, I will only touch on these other roles where they overlap my central theme.

By undergraduate natural and applied science education, I mean not just the education of students specializing in those fields, but education of all students who take any courses in natural and applied sciences. This encompasses a major fraction of the entire undergraduate population. A meaningful exposure to these technical subjects is an increasingly important part of any educated person's background, because of the growing importance of science and technology in modern society. While I suspect that much of what I say applies more broadly than to only undergraduate education in the sciences, I do not have data or expertise to fully substantiate such a claim. Furthermore, successful models for implementing desired changes will depend on the cultures and natures of the disciplines. These are often rather different between science and nonscience departments.

Finally, it is important to delineate the standards by which I judge a "good" or "bad" education in the sciences. My standard is that the programs and individual courses should move the student towards expert competence in the subject. This means acquiring the problem solving approaches and skills, habits of the mind, content knowledge, and beliefs about the nature and relevance of the subject that are like those of practicing experts. In a full program of study, of course, there should be much larger changes than can be achieved in a single course. But even for the nonscience student taking a single science course, the educational goal should be to have them understand science and think about science more like a scientist.

There is a large and growing body of research indicating that post-secondary science education is failing in this regard. Although most of the research has examined students' learning of physics, there is a significant amount of data on the learning of chemistry and some for biology as well. All of these results show a consistent pattern. Most students are learning that the subject is a set of facts that are unrelated to the workings of the world and are simply to be memorized without understanding, and they learn to "solve" science problems by memorizing recipes that are of little use other than passing classroom exams. Furthermore, they are leaving their courses seeing the science as less interesting and relevant than they did when they started. The typical student is *not* learning to see the science like an expert, as a set of interconnected experimentally determined concepts that describe the world. They are also not learning the useful concept-based problem solving methods of experts that can be applied in many different contexts. Below I discuss the reasons for this and how this situation can be changed.

As an aside, I might note that this population of typical students includes most of the future K-12 teachers, so these shortcomings in science education at the post-secondary level have repercussions at all levels of the educational system.

II. Current model for higher education.

The current model of higher education grew in a haphazard unplanned fashion that has left it with traditional practices and modes of organization that in some aspects are poorly matched to modern educational needs.[1] It seems likely that the university grew out of the apprenticeship model of an expert working closely with an apprentice, assigning them challenging tasks and then providing guidance as needed to carry out those tasks, as well as offering ongoing feedback on their work. This model, or its modern day embodiment of "the expert individual tutor," remains the most effective demonstrated approach to education in most areas. As knowledge and population grew, the apprentice model expanded into the university with an increasing number of students for each expert, in order to pass along information more efficiently. The lecture format which still predominates today began long ago, before the invention of the printing press, as an efficient way to pass along information and basic skills such as writing and arithmetic in the absence of written texts. The economies of scale led to this expanding to the current situation of a remote lecturer addressing often several hundreds of largely passive students.

Although it is unclear how effective this model ever was for science education, vast societal and technological changes over the past several decades make it clearly unsuitable for science education today. The most significant of these changes are discussed below.

1) Modern day educational needs and goals are far different from what they were in past centuries or even a few decades ago. The modern economy demands and rewards complex problem solving and communication skills in technical subjects. These are far more important than simple information/knowledge. The new importance of learning complex problem solving skills is frequently at odds with traditional university teaching practices. The lecture model, while conducive to transfer of simple information, lost much of the individualized challenging exercises and feedback that is a critical part of the apprenticeship model for acquiring deep understanding and complex problem solving skills. While this individual instruction was

retained in the British system of tutors for study in sciences, that system is not economically practical for large scale use.

2) Changing student demographics: Until a few decades ago, college education had always been necessary and useful only for a very select elite. Now college has become a basic educational requirement for most occupations in the modern economy, particularly occupations of most importance for general economic growth and personal economic success. This means that a far larger and more diverse fraction of the population is seeking post-secondary education than in the past, and thus a system is needed that can deliver a high quality education to that large diverse population.

It is difficult to adequately emphasize how enormous this demographic change is from the situation that existed when most of our colleges and universities were originally created and their organizational structures established. It is even dramatically different from what existed when many of today's college teachers and administrators were in college. So while many lament how we just need to get back to the "good old days," such lamentations are at best irrelevant and at worst simply ridiculous given today's realities. We face an educational challenge which is unprecedented: the need to effectively teach complex technical knowledge and skills to a large fraction of the total population. The approaches of the past are clearly inadequate to meet this need.

3) Faculty members' responsibilities are far different from what they were several decades ago. This is particularly true at the large research universities that stand at the top of the higher education pyramid and train nearly all the higher education faculty. The modern research university now plays a major role in knowledge acquisition and application in science and engineering, through the efforts of the faculty. Running a research program has become a necessary part of nearly every science and engineering faculty member's activities, and it is often the most well recognized and rewarded part. Such a research program requires the successful faculty member to spend time writing proposals and obtaining research funding, managing graduate students and staff, writing scholarly articles, participating in scholarly societies, and traveling to conferences and lectures. This is much like the demands of running a small (or sometimes not so small) business. Faculty members are also increasingly encouraged by their institutions and governments to take the additional step of converting the knowledge of their research lab into commercial products. This brings additional revenues into the institution and provides highly visible justification for the government expenditures on basic research at universities. When they take this step into commercialization, the faculty members are often literally running a business, in addition to the business-management-like responsibilities of running a university research lab. While good arguments can be made for the value of such faculty driven university research and the creation of spin-off companies, the result is a faculty with new sets of demands and responsibilities that were largely nonexistent at the middle of the last century. These demands must be considered in any discussion of changing higher education.

4) While the above changes are in the educational role and environment of the university, changes of a rather different sort have also taken place; changes in the state of knowledge of how to assess and achieve effective science education. The understanding of how people think and learn, particularly how they learn science, has dramatically improved over the past few decades.[2] While there has never been a shortage of strongly held opinions throughout history as to "better"

educational approaches, now there is a solid and growing body of good research, particularly at the college level in science and engineering, as to what pedagogical approaches work and do not work and with which students and why. There are also empirically established principles about learning emerging from research in educational psychology, cognitive science, and education that provide good theoretical guidance for designing and evaluating educational outcomes and methods. An important part of this research is the better delineation of what makes up expert competence in a technical subject and how this can be more effectively measured.

While there is still much to be learned, there is enormously more known now than existed when the teaching methods in use in most college classrooms today were introduced and standardized. Briefly summarizing a large field, research has established that people do not develop true understanding of a complex subject like science by listening passively to explanations. True understanding only comes through the student actively constructing their own understanding through a process of mentally building on their prior thinking and knowledge through “effortful study”.[3] This construction of learning is dependent on the epistemologies and beliefs they bring to the subject and these are readily affected (positively or negatively) by instructional practices.[4,5] Furthermore, we know that expert competence is made up of several features.[2,3] In addition to factual knowledge, experts have unique mental organizational structures and problem solving skills that facilitate the effective retrieval and useful application of that factual knowledge. Experts also have important metacognitive abilities; they can evaluate and correct their own understanding and thinking processes. The development of these expert "beyond factual" competencies are some of the new ways of thinking that students must construct on their path to “expertness.”

There are important implications of this research for both teaching and assessment. i) The most effective teaching of science is based upon having the student fully mentally engaged with suitably challenging intellectual tasks, determining their thinking, and providing specific targeted and timely feedback on all these relevant facets of their thinking to support the student's ongoing mental construction process. ii) Meaningful assessment of science learning requires tests that are carefully constructed to measure these desired ways of thinking. As such, their design must be based on an understanding of these expert characteristics and how people learn, in addition to a thorough understanding of student thinking about the subject in question. Such assessments go well beyond the simple testing of memorization of facts and problem solving recipes that is the (unintended and unrecognized) function of the typical college examination. Much of the rest of this article concerns how such effective teaching practices and the associated valid assessments of learning can be implemented in the modern university environment.

5) The final dramatic change is in the state of education related technology. Everyone is aware of the enormous increases in the capabilities of information technology (IT) over the past few decades, years, and even months. These offer many fairly obvious opportunities for dramatically changing how teaching is done in colleges and universities, and in the process, making higher education far more effective and more efficient. Unfortunately, these vast opportunities remain largely untapped. While there are a few spectacular examples, generally the educational IT currently available is quite limited in both quantity and quality.

We are now at a watershed in higher education. We are faced with the need for great change, and we have the yet unrealized opportunities for achieving great change. The full use of the research on teaching and learning, particularly as implemented via modern IT, can transform higher education, and allow it to do a far better job of meeting the higher education needs of a modern society. Below I will discuss the characteristics of this hypothetical transformed university, followed by discussion of how to achieve such a transformation.

III. Characteristics of the *optimized university* with some comparisons with the current situation.

While one might envision an ideal university that is totally redesigned and has unlimited resources, it is impractical to imagine that such an institution can be created. So instead I will discuss what an *"optimized"* university might look like. This is a university that provides the best undergraduate education possible within certain basic constraints on resources and organizational structures. The constraints are based on the pragmatic assumptions that resources in support of higher education will not dramatically increase and most of the long standing structures such as disciplines and departments will be largely intact, as will the current broader faculty responsibilities. I do not believe that it is possible to avoid these particular constraints. First, there is no indication that dramatically higher levels of resources are forthcoming for public education. Second, where attempts have been made to create universities with dramatically different organizational structures, such as the new University of California campuses without discipline-based departments, they have effectively reverted to largely traditional structures. It is difficult to see how anything else is possible given the complexity and extensive scale of modern natural and applied science and the limitations of the human mind. There must necessarily be some organizational unit at the level of "extent of material that intelligent human (i.e. faculty member) can reasonably master" that serves as the basic unit of educational organization. While there are many new interdisciplinary areas of activity, from a longer-term perspective, these are primarily a continuation of the historical evolution of disciplines to remain appropriately aligned with the directions in which science and engineering is developing. Therefore, while I assume that the labels and orientation of departments will change, there will remain entities on the size scale (intellectual and number of faculty) of departments. These will continue to be the basic organizational structure of the faculty members and the primary educational unit within the university.

While these external features of the current university and my transformed optimized university will look the same, there will be some dramatic differences. Education in the optimized university will focus on the desired student educational *outcomes* and these outcomes will be measured and achieved through a structure of pervasive thoughtful use of both research on learning and information technology. This focus on learning outcomes is in contrast to current practice of focusing on processes, such as number of students taking certain number of courses covering particular list of topics. If properly implemented, this switch from processes to outcomes will ultimately lead to dramatically improved educational results and improvements in educational efficiency.

Another subtle but important difference of the optimized university are the roles of the student and the faculty member in the learning process. Currently the implicit roles are that the faculty member simply transfers their expertise, as if it were bits of information, to the receptive students, much like pouring water from a large jug into a set of small receptive cups. This model is inconsistent with what we know about how people learn science. In the optimized university the role of the faculty will be as "educational designers," utilizing their knowledge of the discipline and how best to learn that discipline to design optimized educational environments, activities, and assessment. Within those environments, students will have the role of effortful constructors of their understanding. A critical part of the educational design will be the ongoing formative assessment, by which the instructor, assisted by technology, will assess each student's development of mastery and ensure suitable targeted feedback and challenges are provided to him or her to optimize their learning.

III. A. Characteristics of the *optimized* university educational environment

The student will first encounter a choice of academic programs, each of which has a clearly delineated set of educational goals. These goals are created collectively by the relevant faculty in consultation with other stakeholders such as industry, educational systems, and government, and will encompass the full set of skills, knowledge, and ways of thinking that are part of an education. Each academic program will then have a series of courses that are carefully aligned and sequenced to progress toward the program goals. Each course will have its own explicit learning goals that identify what students should be able to do at the completion of the course and relate to the program goals. These learning goals will be also be established by a consensus of the department faculty members, and will be maintained, regularly reviewed, and updated in the normal functioning of academic departments.

In each class, the student will encounter pedagogical approaches, materials, and technology all based on careful research and testing. The student's learning will be measured and guided on an ongoing basis using a variety of tools and technology. The development and improvement of these measurement tools will also be seen as a basic departmental responsibility and will reflect the values of the faculty. Faculty teaching evaluations will be linked to these measures of student learning.

The entire educational process will be driven by these clearly established and measured outcomes of student mastery of detailed educational goals. While it will take substantial investment to produce meaningful measures of outcome to make this possible, the knowledge base and technology now exists to make this feasible on a large scale, and the ultimate returns on this investment will be enormous. This is the only way to ensure that good pedagogical methods and environments are replicated and improved upon and poor ones are eliminated.

When a student starts a class in the optimized university, they will first complete a detailed diagnostic examination that accurately determines their preparation/knowledge-state. This will examine their content and conceptual knowledge of the subject and those subjects that the course builds upon, such as mathematics and related science disciplines. This will also diagnose their beliefs and epistemologies about the subject and how it is best learned. Before they have ever seen an instructor, the instructor would have a profile of their strengths and weaknesses, and the computer would have already flagged serious deficiencies. If these deficiencies are widespread, the student will be guided to enroll in a more appropriate course. Where the deficiencies are localized and not severe, the computer would provide the student with feedback and suitable exercises that they must complete to remedy these deficiencies. This will ensure that the course will begin with all students at roughly the same level of knowledge and competence, and the instructor will have an accurate profile of that level. This will make it possible to design learning environments that are well-matched to the population of students; something that currently is very seldom the case.

This initial extensive diagnostic exam will be the first of regular ongoing evaluations throughout the course of the student's thinking and learning. These evaluations will be linked to targeted timely feedback to both student and instructor. Such a scale of evaluation and feedback will only be practical with widespread use of information technology. This evaluation and feedback will be largely provided by online homework systems that include intelligent grading and tutoring programs. In this way, much of the individualized evaluation and feedback that has been lost, in moving from a personal expert tutor model to a single instructor with many students in a class model, will be reestablished. This will be done in an economically practical manner by using IT to greatly extend the instructor's capabilities.

III. B. Educational technology

The type of technology required for these purposes has been demonstrated in certain specific areas and has been shown to be highly effective under limited experimental conditions, and in a few of cases, fairly large-scale experiments. However, it is used in an extremely limited fashion in education. The quantity and quality of what exists has barely scratched the surface of what is needed and what could readily be created, if there was support to do so. For it to be created there must be a viable business model (which does not currently exist) driving its development by private industry, or governmental support. There are major problems with the creation of such a viable model, as long as there is no link between educational outcomes and resources, as discussed below. These are much the same factors that result in textbooks that blatantly conflict with well-established pedagogical principles. Assuming resources can be found to carry out the development of these valuable educational technologies, their development must be guided by knowledge of the specific disciplines and research on how people learn. A clear understanding of the educational capabilities and limitations of IT and careful testing of the products are also essential.

The college classroom is primarily precomputer in its level of technology use. However there are many new educational technologies that have been demonstrated to be highly effective and will be used widely in the optimized university. These new technologies have the capability to transform the higher education system, in much the way the high technology industrial setting has been transformed from what it was in the 1960s. The educational applications of IT waiting to be fully realized range from the mundane but time (and hence money) saving to highly sophisticated new methods for learning. It would require a far longer paper to do justice to this subject, but some of these applications include technology for new teaching methods (interactive simulations, intelligent tutors, sophisticated diagnostic capabilities, student in-class personal response systems "clickers"), improved class organization and management systems, archival systems for educational materials and data, and new modes of presenting material and enhancing communication by linking students with each other and faculty.

III. C. Research based instruction.

The faculty of the optimized university will have sophisticated "pedagogical content knowledge",[6] in addition to the usual content knowledge for every course they teach. This "pedagogical content knowledge" means knowing: how the content and skills are best learned, what common student difficulties are encountered in learning it, what approaches are most effective in helping students overcome those difficulties, and how best to motivate students to master the subject. What is required is knowledge of the relevant research on learning, and assessment of learning, as it specifically applies to the subject in question. In the optimized university, a general knowledge about how people learn science will be part of every faculty members' basic competence, and the many subject specific pedagogical and assessment issues will be fully researched, and detailed information on them will be readily available to every faculty member. When a faculty member is starting out to teach a course, their first step would be to study these course specific pedagogical content materials.

While the student will likely still see an instructor in charge of each particular course in the optimized university, the relationship between the course and instructor will be rather different. The department will have the basic responsibility for each course and what students are learning in it. The instructor of a course will thus be working as part of a collective enterprise to optimize the course within the goals, guidelines, and assessments established by the department. Faculty members will work in teams to first establish clear educational goals from large scale to specific topic level, and then collectively develop and refine approaches, materials, and assessment tools. The products produced by these collective efforts will be routinely reused and improved upon to provide ever more efficient and effective instruction. Members of the faculty team will each share their strengths to achieve a whole that is greater than the sum of its parts, and, in the process, expertise will be shared so that younger, less experienced, faculty will rapidly gain teaching expertise. This is in stark contrast with the current system where teaching is an isolated activity in which faculty set their own agendas and goals for the courses they teach, and they struggle in isolation to teach the subject effectively. Although the collaborative approach described above is highly unusual in teaching, it is not unprecedented. Also, all of these activities and modes of operating are the norm in the modern scientific research lab. Hence the problem is not one of convincing faculty to function in a radically new manner, but rather the lesser challenge of getting them to see how approaches that they know and recognize as very effective in one setting (the research lab) can be equally effective in another (teaching).

The effect of teaching and learning styles

A common claim is that such a collective approach to teaching would fail because what are effective or ineffective teaching and learning styles are totally or largely dependent on the individual personalities of the teacher and learner. I would argue that such a claim is quite inconsistent with a large amount of research data. All normal human brains function in the same basic way, and research has clearly established that there are very general features of effective teaching and learning. While there of course are individual distinctions, particularly in the learners, these distinctions are small compared to the range of teaching approaches for which there are advocates. For example, there is extensive physics education research literature examining the effectiveness of various teaching practices. This consistently shows that practices that increase the average learning for a class also increase the learning for each of the subgroups of low, medium, and high achieving students in the class. The individual student distinctions with respect to effective teaching styles are evident primarily only at the much finer level of the student thinking on specific topics. Thus, they are best addressed by the careful evaluation of thinking and providing appropriate feedback as described above, rather than trying a wide variety of teaching approaches in the hope that what fails for some students might be successful for others.

The myths of the innate super teacher and teaching as an individualized "art form":

Similarly, it is often claimed that teaching effectiveness is dominated by the personality of the instructor; some "have it" and others do not (with the implicit assumption that those who don't have it can't be good teachers, no matter how hard they try). A frequent corollary is that what is or is not an effective teaching method depends largely on the personality of the instructor. These claims are also clearly contradicted by the data, again from Physics Education research. This shows that when careful assessment of student learning is carried out, even the most dynamic, interesting, and entertaining lecturers fail to achieve good learning outcomes with the traditional lecture format, whereas “ordinary” teachers can achieve much better learning outcomes if they implement research-proven effective practices.

Since this issue of individual teaching styles for faculty and individual learning styles for students is such an important and often misunderstood point, let me give more detail. I would argue that at a basic level, the same factors are important for learning for everyone. All people learn best by actively trying to understand something and receiving well-targeted feedback to guide their thinking. They all suffer from the limitations of short and long term memory and the related demands and limitations of cognitive load. Everyone learns a subject better if they are motivated to learn, including understanding why it is of value to them to learn it, etc. These similarities dwarf the differences. While few of these claims have been exhaustively tested, there are a number of specific practices that benefit nearly everyone on which they have been tested, such as having timely well-targeted feedback that directly addresses one’s reasoning and says what is right and wrong about it. While there are frequent claims about students having very different learning styles, the research support for this involves studies of tasks such as memorizing lists of numbers and whether this can be done better when the numbers are spoken or written. Such data is of very questionable relevance to the learning of science. Individual differences in motivation due to different backgrounds and experiences are clearly evident, but these are much like differences in preparation, and are learned rather than being innate. Similarly, differences are evident when one gets down to the level of student thinking on specific topics.

However, often even these differences are relatively small. The thinking of individual students has variations, but they tend to be fairly limited in extent. Detailed research into both students’ general beliefs about science and learning science and the learning of specific topics reveals that there are some rather general shared characteristics. As an example of the latter, in teaching students to understand basic electricity, nearly all students who do not understand it will have one of a few incorrect ways of thinking about it. The misconception that electric current gets “used up” in flowing through a light bulb is an example of a widely shared misconception. Knowing these categories, being able to recognize when a student’s thinking falls into one of them, and knowing the most effective feedback to help them correct their thinking is an example of the “pedagogical content knowledge” that a good teacher will have. If a teacher has this knowledge, they are much more likely to be effective than if they do not, quite independent of their personality or that of the student.

The role of the computer as tutor

In those many subject areas where science education research can establish a few common difficulties, computer technology can take over much of the role of an individual expert tutor. The computer asks questions that are known to delineate the category of thinking of the student. Then, based on the student’s responses to those questions, the computer provides the feedback that research has shown best helps them to correct their particular incorrect thinking and/or reinforces their correct thinking. Thus, one can provide most of the benefits of an expert individual tutor, but in a manner that is economically practical for very widespread use. This feedback step requires software that, while clearly practical with today’s technology and understanding of learning, and even available for a few very specific topics,[9,10] is more complex and more difficult to develop than a program that simply identifies student thinking. Thus, the intermediate step on the way to the fully optimized university would be to have software that could quickly determine the student thinking and misunderstandings and provide the instructor with that knowledge. The instructor could then adjust instruction to provide immediate and helpful feedback to the students. This is quite similar to what already exists in the rapidly growing number of classrooms that use student personal response systems ("clickers"), where the students are asked questions in class that probe their thinking and misunderstandings. The students respond to the questions using their clickers, and a computer records their responses and displays it to the instructor in real time.

It is sometimes mistakenly thought that the use of intelligent tutoring systems to teach science is quite mechanistic, and as such is either impossible, or if it is possible, it will totally replace the instructor with a computer program. For learning relatively routine low-level skills and knowledge, the latter probably is true. For example, both faculty and students characterize learning introductory anatomy as being primarily about memorizing names, locations, and functions in the body. A good computer program could accomplish this, and determine that it had been mastered, better and far less expensively than a human instructor. However, for most of the more complex scientific information and skills, the value of the instructor remains, and I would argue it is enhanced rather than diminished by the use of technology. In my own educational R&D efforts, as well as those of others, incorporating the various research based approaches, particularly those addressing student beliefs, motivation, engagement, and understanding, results in nearly all students becoming far more engaged in the subject. Students who are engaged in the subject explore it in more depth and examine how the ideas apply in much wider range of contexts. The result is far more numerous and deeper questions, so an instructor who is an expert in the subject is essential-- considerably more so than in the case of traditional science teaching. Rather than merely a filterer and transmitter of information, the instructor is now routinely called upon to help students examine and understand the ideas at a much higher cognitive level. Also, via the technology, this instructor will be far better informed about the students' strengths and weaknesses and thus can have much more educationally effective interactions with them.

Creation of valid assessment tools

I am frequently asked what the difference is between the kind of assessments I am calling for in the discussions above and the usual examinations that are used in college classes. The flippant answer is, "at least six months of hard work." When the typical science exam is examined carefully, in spite of the best but usually untrained efforts of instructors, most students can and do complete them successfully using strategies based on simple memorization of facts and problem solving recipes. Valid assessments of the desired deeper understanding require a detailed examination of student thinking in the context of the specific subject material and the specific understanding and problem solving skills that are the goals of the course. Only then is it possible to create an assessment instrument that provides the requisite probing of student thinking. The experience of assessment experts show that even exams constructed on such a foundation must still be carefully tested for validity and reliability with students, before it is possible to be confident of their value.[7] Some examples of such assessment instruments that have been created for use in introductory physics include the Force Concepts Inventory (FCI), the Force and Motion Concept Exam (FMCE), the Basic Electricity and Magnetism Assessment (BEMA), and the Colorado Learning Attitudes about Science Survey (CLASS).[8] A similar design concept but a more extensive development effort is required when the assessment device is part of a software system that provides real time feedback to guide learning. Examples of such systems are the Diagnoser program of Minstrell[9] for diagnosing student understanding and misunderstanding in areas of high school physics, and the Cognitive Tutoring systems of Koedinger and coworkers for teaching algebra.[10]

Class size.

In considering how to optimize higher education, it is impossible to avoid the question of optimum class sizes. While everyone involved would prefer individualized instruction with class sizes of one or two, this is clearly impractical. From a purely economic point of view, the larger the class sizes the better. So the real question is, what is the tradeoff between class size and learning that is the optimum use of resources? An extension of this is, do we even need classes anymore? Can't we just teach everything online with the proper software? I would argue that while online classes could easily replace classes involving students sitting in a cavernous auditorium listening passively to a lecture, it is much harder to see how they can replace classes designed around the interactive engagement of students in ways that have been found to be much more educationally effective (the norm in my optimized university). In these sorts of classes, there are social interactions (student-student and student-instructor discussions) that clearly play a large educational role.

These same teaching style issues are relevant to the question of optimum class size. A class that relies on the traditional passive lecture format is equally ineffective with 20 or 200 students. Also, a large fraction of the learning in most good science courses happens outside of the classroom, and this outside-of-classroom learning is only indirectly affected by class size. However, without the use of technology it is clearly more difficult to achieve pedagogically effective social interactions and targeted individual feedback in a class of 200 students than in a class of 20. So the uses of research and technology as discussed above to make classes more intellectually engaging and educationally effective often have the most obvious gains for large lecture courses. There are demonstrations of classes of 200 or more achieving very good learning gains by utilizing technology and research based practices such as: clickers and peer instruction, good computer graded homework systems, encouragement of pedagogically effective student-student collaboration, extensive course webpages, and email and online communications and survey systems. Learning gains in such classes can be as good as the best achieved in much smaller classes. Therefore, I do not think it is possible yet to say what class size would result in the optimization of learning within a fixed amount of resources, and the standard mantra of "smaller is better" is almost certainly not the optimum. I have searched for data on this subject and have found very little. The very limited data I know about (much of which comes from work of my group and collaborators) suggests that the optimum depends on room layout, and probably other factors, and is less than 400 but is perhaps more than 50. It is clearly in the interests of higher education to carry out studies on the tradeoffs between learning and class size.

III. D. Gains in effectiveness and efficiency.

The optimized university will have enormous improvements in effectiveness and efficiency. The effectiveness comes from the value of having research-tested teaching methods in widespread use by faculty who understand and know how to use them; by the use of extensive technology-based formative assessment so that each student is being challenged at the level where they can successfully build their understanding and expertise at the optimum rate; and by the timely and targeted feedback provided by technology-aided instruction to guide the student's thinking. Also, students will be receiving instruction that reflects an ever improving state-of-the-art knowledge as to what topics and skills are important in the subject in question and how to help students best learn these. This ongoing improvement will take place because of the existence of good outcome assessments and the widespread dissemination of results and dissemination and duplication of successful practices.

The potential gains in efficiency can be seen by comparing the operation of the optimized university with the typical current university, where there is a largely unplanned and ill-structured course of instruction based on tradition, textbooks, or habit. Here is a list of the major gains in efficiency.

<u>1. Disseminating and copying what works.</u>

Each time a faculty member now goes to teach a new course, they typically reinvent it from the beginning. Thus they must spend time creating new learning goals, lectures, exams, etc., and at the end of all this large amount of labor, the result is basically the product of an inexperienced amateur. After some iterations teaching the class, the good teacher will gain a better understanding of what does and does not work, and there will be improvements, but these will be never exceed the knowledge, experience, and capabilities of that particular faculty member. Sadly, as soon as a new faculty member takes over the course, the situation reverts to the beginning. Quite aside from the questionable educational effectiveness of this approach, it uses up a great deal of faculty time by forcing each to redo the work of their predecessors, as well as often repeating their mistakes.

It is often said that teaching is an individual activity that each person must do in their own style and therefore such reinvention and abandonment of what has been done in the past, with its corresponding inefficiencies and deficiencies, is inherent to the teaching process. However, this argument could be made with just as much validity (or lack thereof) to the scientific research that these same faculty members are engaged in. In science research, it is obvious how it is possible for researchers to continue to build upon and extend the advances of their predecessors with their own quite individual efforts and styles. Through this process they achieve results far beyond the capabilities of any single person. There is no reason why the teaching of science cannot be as successful as the practice of science in this regard.

<u>2. Eliminating the problem of vast discrepancies in student backgrounds.</u>

The greatest source of inefficiency in the current system of higher education is the enormous variations in student backgrounds (knowledge, skills, beliefs about how to learn and why to learn) encountered in nearly every undergraduate science course. This variation in students, combined with the lack of good ways to measure and respond to those differences, causes great difficulties and wasted time for faculty and students alike. The typical college science class, when it is going quite well, has perhaps 30% of the class bored because they already mastered the material (often in a previous course), 30% of the class so lost they are not learning anything (often because of a small but crucial deficiency such as knowledge of a particular terminology or mathematical technique), and 40% are getting some educational value. This 40% that are learning something is probably the best case scenario; often, because of lack of knowledge about the students or pedagogical miscalculation by the teacher, that fraction is much less. This means that a large fraction of both student and faculty time is being wasted because there is not a good way to routinely assess student learning. If there were, all students could achieve a clearly established mastery and the faculty could align the level and material covered accordingly. The resulting smaller spreads in student preparation, and better knowledge of that preparation by the instructor, would improve efficiency by making it possible to design courses that are optimized for the learning of the great majority of students, rather than the current inefficient compromise that is not well suited for anyone.

The extreme case of variation in backgrounds is illustrated by those students who are enrolled in a course for which their preparation is clearly inadequate. This is a frequent source of anguish and large amounts of wasted time for both students and faculty. The wasted time and hardship that students encounter in trying to master material, when their inadequate preparation makes it impossible, is obvious. What is not so obvious is the large hidden cost in instructor's time spent dealing with students who are not adequately prepared. These students often take up a disproportionate amount of instructional time, both in the need to provide them with extra assistance, and in dealing with the repercussions of failing students-- complaints from students and parents, pleas and arguments for regrading, special exemptions, etc. In teaching a class of 100 students or more, a conscientious faculty member will spend a significant amount of time in such activities, time that comes out of what is available for the education of the properly prepared student. This inefficiency could be easily avoided if one has good diagnostic exams, as discussed above, to assess student preparation and ensure students are only enrolled in classes for which they are adequately prepared. This would benefit both the students who are not prepared for the course, and the students who are properly prepared.

<u>3. Avoiding unnecessary repetition.</u>

Another striking example of inefficiency of the current system is the way in which the same science topics are covered repeatedly in the curriculum for a science major, but each time covered so rapidly that students do not achieve mastery. This sort of repetition has been noted as one of the distinctions between the K-12 education system in the US and several Asian countries that score far higher on mastery of science and math in international comparisons.[11] As an example from higher education, an undergraduate physics major will cover nearly every specific topic two to three times over their course of study. Other sciences have similar repetition in their curricula. I am aware of no evidence indicating value to the current system of rapidly covering the same material multiple times. On the contrary, research shows that when students develop misconceptions from their initial instruction, these tend to be maintained throughout subsequent instruction. On the other hand, when they have true mastery and understanding of the topic, it is robust and sustained. Thus, it is likely that such repetition of coverage is not only unnecessary but is even detrimental. Careful measurement of student learning to ensure they master the topic when it is first encountered will make it possible to design curriculum that avoids repeating coverage of the same material. This will eliminate an enormous inefficiency in the current system.

<u>4. Eliminating expensive faculty time being spent on low-level tasks.</u>

Another easily remedied inefficiency in the current system is the large amount of faculty "teaching" time spent on rather low-level tasks that could be performed by far less expert and lower cost staff. This involves routine class maintenance, recording of grades, dealing with students who are dropping or adding classes, dealing with special student circumstances such as missing assignments or exams due to medical or family emergencies, etc. The fraction of the "teaching" time required for dealing with these issues scales with the number of students. When typical class sizes were small this burden was insignificant, and the common organizational structure of having faculty handling such tasks developed during an era of such small class sizes. As economies of scale have driven up class sizes, there have seldom been appropriate organizational changes in response. As a result, for large classes these low level tasks can take up a large amount of faculty time, because there is seldom if ever support for staff to carry out these tasks.

Some, although not all, of this time required for general class management could be handled by technology. There has been substantial progress in course management software, but, surprisingly, commercial products are still far from optimum. There are many individual efforts to develop better software, but the scales are such that these systems can never be as well maintained and easy to use as a large-scale commercial product. There are market reasons why commercial development has lagged in this area, but this is clearly an area where suitable investment would provide substantial returns. The optimized university will have suitable software and staff to avoid using any expensive faculty time on such tasks.

5. Optimizing the cost and effectiveness of support and feedback to students.

In the optimized university, all the faculty instructional time will be spent on high-level educational tasks befitting both their expertise and cost, such as: delineation of desired expert skills in the discipline, pedagogical design and testing including enhancement of previous work, high intellectual level student interaction, and guidance of TAs. There will be a fairly clear hierarchy of support for student learning that will provide the optimum benefit for a given amount of financial resources. This is not a hypothetical model; I have implemented in several science courses.

In this model the student is first mentally engaged by being given some suitable intellectual challenge, most commonly a homework problem carefully designed by a faculty member. In the current system, a student will typically work on homework in isolation (encouraged in this relatively unproductive activity by general policies and curve-based competitive grading systems). If they receive any feedback to guide them in their learning, it will likely come in the form of submitting a solution, and one or two weeks later finding out if their answer was correct. Research shows that such feedback serves very little if any pedagogical function. If they have a small class and a dedicated teacher, they may get more useful feedback by talking to the instructor about the problem. However, this is seldom practical on a widespread scale for large classes. Also, it is often a poor use of resources, since frequently the feedback, although quite necessary for the student to make progress, is very low level ("The reason your answer did not make sense was not because you misunderstood the concept, but because when you put this number into your calculator, you accidentally put in 200 instead of 2000.")

In contrast, in the optimized university, the student will have many levels of support and feedback. At the lowest level they will have intelligent tutoring systems and/or collaborative fellow students (in person or online) providing them with feedback. So, rather than struggling in isolation with the problem and making little progress for hours, they will have the fellow student quickly point out to them their calculator error, or nearly as often, they will discover their own error in the process of explaining to their fellow student how they are trying to do the problem. Structures and grading policies of the course will encourage such student-student collaboration and their associated well-established pedagogical value. When the difficulties become so great that collaboration and feedback from fellow students is not sufficient to allow the student to make further progress, then there will be trained undergraduate and then graduate teaching assistants to provide the necessary guidance. There is a large range of student difficulties and situations where such guidance is as adequate, and even sometimes superior, to that provided by a faculty member, because the teaching assistants can have a better perspective on the student's thinking and thus provide more effective feedback. Finally, for the most challenging issues that demand expertise beyond that of the TAs, the faculty will provide the necessary feedback. Such a hierarchical support system can work by taking advantage of the capabilities of modern communication systems. The far more expensive faculty time is then utilized when, and only when, it is required. This model allows one to provide a highly supportive and effective educational environment for large numbers of students, at a reasonable cost.

6. Training Teaching Assistants to become important contributors to undergraduate education.
Graduate student teaching assistants (TAs) have taken on an increasingly large part of the teaching at research universities. While this clearly has economic benefits because TAs are far less expensive than regular faculty, it is a source of frequent and usually well justified complaints. As in so much of the higher education system, the use of teaching assistants developed in a haphazard way and hence could be dramatically improved by some strategic planning and optimization. Originally, TAs were used for routine grading of exams and homework. Then economic pressures moved them into lower level teaching jobs such as overseeing students in labs, where relatively little supervision or planning was required. TAs can now be found carrying out a large fraction of the teaching in many situations. Just as there is little or no attention to training faculty for teaching – because there has long been the implicit, though now thoroughly discredited, assumption that if one masters the content, one can teach it effectively – a similar assumption has been made about teaching assistants. This has resulted in many classes being staffed with poorly trained and poorly supervised TAs whose teaching results in many calls to replace TAs in the classroom with faculty. However, this does not make sense either economically or educationally. There are now clearly proven examples of how well designed and tested training programs can routinely produce extremely well-qualified TAs who provide excellent educational experiences for undergraduates by every measure, and for some (though not all) aspects, better than a faculty member.[12] Such TA training programs do require small investments (several days of time for each TA, plus faculty oversight), but the return on this investment has been clearly demonstrated to be extremely high in terms of educational value and student satisfaction. From the perspective of optimizing resources to provide the best possible undergraduate education at a reasonable cost, well-trained TAs with suitable faculty supervision clearly makes sense.

7. Optimizing the effectiveness and reducing the costs of teaching laboratories.
A unique aspect of science instruction where there is a great deal of inefficiency is the teaching laboratory. Undergraduate teaching laboratories are particularly expensive in terms of facilities and student and faculty time, and, as they typically function, are doing a poor job of achieving the desired educational goals. This is a subject that attracts particular passion. Most faculty members feel strongly that because experimental research is such an essential part of doing science, laboratory classes must be an equally essential part of science instruction. The argument usually given is that "This is the only way for students to learn how science is actually done!" However, the educational research reveals no indication that the typical laboratory class actually achieves this pedagogical function, and considerable evidence that it does not.[13] This educational

failure arises generally from rather poorly thought out and conflicting educational goals for the lab classes and a dramatic mismatch between faculty intentions and pervasive student perceptions and cognitive practices in lab courses. Considerable improvement in effectiveness and efficiency could be achieved by a judicious examination of the educational goals of science lab classes, the assessment of how well they are achieving those goals, and the best and most cost effective ways to reach those goals.

III. E. Balance of research and teaching
-- evaluation and rewards.

In the modern university, there is an ongoing question as to the appropriate balance of teaching and research, and this remains a question for my optimized university. There is no clear answer to this question. Both teaching and research are essential components of the modern research university and are vital contributions to society. It would be unwise to abandon either. As discussed in the previous section, without changing the current balance of faculty time between research and teaching, there is an enormous potential for improvement in educational effectiveness and efficiency. So I would argue that the best approach would be to achieve those improvements and examine the results, before considering changing the current balance. Also, it is hard to imagine that a faculty member could teach expert competence in an area of modern science and technology unless they have been active in the field themselves for much of their careers. The complexity and rapidity of progress in these fields today are such that faculty simply cannot remain sufficiently expert in the subjects in which they are educating students, if they must rely on teaching the subject only based on what they themselves learned in school. Thus maintaining an active research program in a department clearly serves to enhance the desired faculty expertise.

While research universities currently reward research more than teaching, there is little alternative since the measures of achievement in research are so much better than the measures for teaching. There currently exist very few valid measures of what students are actually learning in any particular course. This makes it impossible evaluate and hence reward actual teaching achievement. There are student evaluations, but these primarily measure the popularity of the faculty member and are readily swayed by a number of factors that are unrelated (and sometimes anti-correlated) with student learning. A few valid assessments have been created for university physics courses as mentioned above, and these have now provided compelling evidence that a faculty member can receive very high student evaluations while the students are consistently learning very little. As long as there is no valid way to generally evaluate teaching in terms of student learning, it is pointless to discuss whether or not teaching is appropriately valued and rewarded. However, when good measures of student learning outcomes are available, this will dramatically change. Then universities will have the capability to suitably evaluate and reward good teaching. Only then will it be possible to make rational decisions as to the appropriate weighting of research and teaching in higher education.

Balance of Research and Teaching
-- role of the authentic research experience in undergraduate science education.

Another aspect of the discussion of this balance is the great educational value to students in having an authentic research experience (ARE) as part of their education. By ARE I mean a research experience that is actually creating new knowledge, and hence involves all the challenges this entails for dealing with real world constraints, requirements for establishing validity of results, and communicating and justifying results according to the standards of the scientific community. This is always assumed to be an essential part of graduate education, but there is also great value for undergraduates to have such experiences, even if they are necessarily limited. Although at present ARE has typically not been integrated into the undergraduate curriculum, the educational value of authentic undergraduate research experience is well recognized by potential employers. When a company is considering hiring a fresh bachelor's degree recipient in physics, for example, they are usually far more interested in what undergraduate research experiences the student may have had and how they performed in them than what courses the student completed. This is equally true when the student is being considered for admission to physics graduate school and, after they are in graduate school, when they are being considered for a graduate research assistant position. Ironically, even though experts in the field (including faculty members) clearly demonstrate their recognition of the value of such experiences relative to coursework, those same faculty members often ignore this value when they set the undergraduate curriculum. At most institutions, undergraduate research experience is relegated to being a useful but nonessential "add on" to the required program of coursework. Also, there is little or no thought given to designing courses that will best prepare undergraduate students to operate effectively in the real research environment. Finally, supervising undergraduate students in research is typically not recognized as an official part of a faculty member's teaching load at either the departmental or the institutional level.

Although many undergraduates in the pure and applied sciences manage to get some research experience in spite of these obstacles, it is clear that this could be a far larger fraction if it was treated as a serious part of the undergraduate education. Having a sequence of courses designed to provide the skills required to function in the research environment would dramatically cut down on the large amounts of one-on-one faculty instruction required to bring a student up to a basic level of competence in the research lab. This would then allow faculty to supervise significantly more undergraduate research students. Counting this supervision as part of each faculty members' formal teaching load and readjusting teaching loads accordingly would further increase the capacity to provide students with ARE.

A few smaller institutions (e.g. Reed College, Oberlin) have such organizational arrangements and curriculum design to make a research experience part of every undergraduate science student's education. In the optimized university of the future, ARE and preparation for it will be a regular part of the curriculum. In contrast to skills gained in most classes, the skills required to participate effectively in a research lab, including communication skills, are a much closer match to skills needed in a technical occupation. Thus, such a curriculum will also better prepare students for the workplace. Thus such a curriculum would be of value, even in situations where there are not enough faculty and laboratories to provide all undergraduates with ARE.

In summary, the great educational value of undergraduate students participating in authentic research also needs to be recognized in the balancing of research and teaching. By optimizing the system, one can preserve the current level of time and productivity in research for the typical faculty member, while also providing the educational benefits of authentic research to a much larger fraction of the undergraduates.

IV. Impediments to optimizing the university.

There are some substantial impediments to moving from the current situation to the optimized university. One is simply the ever present human inertia against breaking with tradition. A larger difficulty is that University governing systems are poorly matched to making changes on a time scale that is rapid relative to the faculty lifespan (several decades). The tendency towards rather short-lived upper administration, particularly in the US where the tenure of public university presidents now averages less than 5 years, combined with shared faculty governance with a faculty that have careers lasting decades, effectively puts the administration in a very weak leadership position. In this regard, there does seem to be some difference between US and BC universities. In the US, governing boards and the position of public university president has become highly political and very subject to the whims of public events, success of college athletic teams, and political intrigue, thereby greatly weakening and distracting the administrative academic leadership. Unfortunately, at the same time that administrative leadership is being weakened, the size and complexity of modern research universities in both the US and Canada has grown to be too great for regular faculty to be sufficiently well-informed and experienced to make major institutional policy decisions. Faculty members simply do not have the time to become sufficiently aware of all the issues and pressures, but they remain a powerful entrenched body that can hinder change. This combination of factors reduces the organizational capacity to carry out useful long-term strategic planning, investment, and implementation to achieve desired changes, such as the optimization of undergraduate education described above. In the US, these factors have arguably nearly paralyzed the ability of public universities to carry out strategic change. It would be wise for BC to heed and continue to avoid the pitfalls of short-lived politicized governing boards and administrations and the distorting influences of large intercollegiate athletic programs. The painful lessons from the US are all too obvious.

Another complication to realizing change is that that the actual “ownership” of educational activities rests almost solely within departments. Realistically, this is necessary. It is impossible, for example, for someone with a background in history, or even in a science such as physics, to be able to say what students should be learning in their biology classes. However, this also means that educational change must happen at the departmental level-- it is very difficult to mandate it from a higher level and achieve the desired effect. Thus educational reform efforts almost certainly have to be based on a model for change at the departmental level.

The final, and arguably the most serious hurdle to systemic improvement in undergraduate education is the lack of a suitable financial model for supporting the necessary educational research and development. Much is required for the optimized university. Student learning goals need to be established that span the desired knowledge, conceptual understanding, skills, and expert beliefs. Then the instructional activities need to be designed and assessed based on these goals. Similarly, the technological tools needed must be created and tested. Although the ultimate value of such efforts is clear, and it is largely a one-time investment, there is a lack of support for such an investment. This lack is fundamentally the result of the current lack of coupling between resources and educational outcomes. There is no connection between the educational value provided by an institution of higher education and the support of that institution, or the support of the faculty who provide that educational value. Hence, there are no incentives for educational change built into the system, and there are several disincentives. There are a number of reasons for this lack of connection between support and outcomes, but I suspect that the most important one is the current lack of adequate measures of educational added-value provided by courses and programs. To repeat myself, meaningful evaluation of educational outcomes in the learning of science is essentially nonexistent and is an essential first step for major improvement.

V. The departmental-level, incentive-driven model for change.

I discussed the many impediments to change in universities, but it is useful to examine the one example of a large and rapid change that has occurred fairly recently within universities. This provides useful lessons for how to accomplish other types of change. The change that I am referring to was the enormous growth in the research enterprise that is now such an important part of large universities and provides a major service to society. The three key factors in this change were 1) it was faculty driven, 2) there were clear measures of success, and 3) there were clear incentives at both the level of the individual faculty member and the department. This change was brought about by faculty seeing the incentives of external research funding that would allow them to do more science, which would in turn increase their status both locally and among the wider community of scientists in their discipline and allow them to contribute to society in new and important ways. Transformation happened at the departmental-wide level, because departments primarily determine faculty hiring, review, and salaries; so the values of the department fuels or inhibits faculty change. There were clear incentives to departments to encourage faculty research activities (increased funding, larger and better facilities, increased prestige, better students) and there were clear measures of outcomes (research dollars brought in, papers published, work cited, scientific awards, departmental rankings) that became collectively accepted. These outcome measures were reflected in the departmental and institutional level evaluation and reward systems, and in turn drove the job market for faculty members that had a high level of success by these measures. This in turn resulted in market forces that impact other colleges and universities; to hire the best faculty it was necessary for an institution to encourage and support research activities. The result was a major transformation of universities that was largely brought about by entrepreneurial faculty who saw clear incentives for their efforts. While support and encouragement from the higher administration was important, the change was carried out at the faculty member and department level.

This example can serve as a model for how to take advantage of opportunities to improve undergraduate education in the sciences. Following this model, one needs to put in place the same characteristics that drove the growth of the science research activities at universities. That means having incentives such as additional funding for individual faculty members and departments that are linked to clear measures of educational outcomes. As such measures are established and so a faculty member can judge their educational efforts accordingly, the nonfinancial rewards such as the prestige that comes along with publications and invited lectures will likely happen automatically. There already are venues for scholarly publications and talks on science pedagogy, and faculty members will be eager to show off their educational accomplishments as measured by the community standards. There are some indications of this sort of progression already beginning to taking place in undergraduate physics education.

Research on physics education at the undergraduate level is emerging as a legitimate form of research in physics departments. Out of that research have come a number of insights on student learning and well developed tools for assessing particular aspects of learning, one of which is the Force Concepts Inventory (FCI) that tests a student's mastery of the basic concepts of force and motion covered in every first semester physics class. The FCI test has now become a community standard and has been given in many physics classes around the world. Its very clear data showing both the failure of traditional lecture-based instruction and the superiority of some other approaches in achieving student mastery has driven transformation of introductory physics teaching at many institutions of higher education. It has encouraged instructors to experiment with different instructional methods and, when they showed improvement in FCI gains, they proudly present their work in publications and talks. This results in corresponding recognition and replication. Out of this work has come the demonstration of approaches that achieve much better results on the FCI test, some of which are inexpensive and relatively easy to implement. These approaches have had the biggest impact on instructional practices. They include such methods as: use of Peer Instruction (posing questions to students in class and having them discuss), particularly when facilitated by personal response systems; and replacing traditional recitations with carefully tested group problem solving exercises such as the "tutorials" developed by the University of Washington physics education research group. The faculty members who developed these methods are now receiving frequent requests to speak, prestigious journals such as Science, Nature, and Physics Today are starting to solicit and publish articles on science education, and the American Physical Society is adding a new award for contributions in the area of physics education. It seems likely that the underlying characteristics that have made this model of change work in introductory physics education would work as well in sciences more broadly.

An important aspect to this model is the manner in which it treats science education as a science. It is well established that one of the major difficulties of organizational change is that the self-identities of the professionals involved are threatened, and this engenders great resistance. In redefining the role of the instructor, as discussed above, such reactions certainly will occur. This is particularly true for those instructors whose identity as a teacher is based on the self-image of a sage dispensing wisdom to an appreciative audience of students. In many work practice innovations, it has been shown[14] that the success of a change process often depends heavily on how it is related to culturally based practices of the organization and impact on core members' identities and sense of self. As discussed by Rogers[15], what seems to matter most in individuals' attitudes and responses to proposed innovation is the way in which they perceive the relative value of any change -- that is, whether and how they can link what is proposed to what they already value. There are two rather distinct aspects of the culture of a science department at a research university, the culture of teaching and the culture of scientific research. The model of the optimized university requires changing the teaching culture, but the change relies heavily on the values and practices of the research culture. This shift will bring the teaching and research aspects of the culture much closer, which should facilitate the change process. Faculty members who do scientific research understand and value quantitative results and will adjust their beliefs if presented with convincing quantitative data on student learning outcomes. Also, they understand and value conceptual and higher order thinking skills and "expert attitudes" about science. My basic hypothesis is that, because of these characteristics, most science faculty will change how they approach teaching if: i) they are shown meaningful ways to assess student learning, particularly higher order thinking skills, and ii) they see that these assessments quantitatively demonstrate the superiority of new research-based and technology supported teaching methods over traditional approaches. Essentially, this model will have the self-identity of faculty members as scientists expand to include their identities as teachers of science. However, this requires that their teaching practices and measures of success are based on research, empirically grounded principles, and objective data.

Once the faculty members and their departments are committed to transforming and improving undergraduate education, there are still three significant hurdles. First, typical science faculty members have little knowledge of research on learning, meaningful assessment, and effective research-based teaching practices. Second, they do not have time to go out and learn about these things on their own, let alone put them into practice effectively in actual courses while maintaining their current level of other responsibilities for research and service. Third, most do not have knowledgeable, interested colleagues with whom they can discuss and develop these novel teaching ideas. The importance of having such an interested and involved community of scholars for exploring new ideas and laying out new directions is readily apparent though seldom discussed when considering successes in the science disciplines.

There are a variety of models one could consider for how to overcome these hurdles. Here I will discuss one specific model that I have developed and demonstrated successfully on a small scale. This involves the use of science education specialists. These specialists are typically fresh Ph.D.s in the respective science discipline who are interested in careers in teaching and/or science education research. They are hired by the department, but are given intensive training (by me and/or my associates) in science education research. (Experience has shown that it is usually easier to find Ph.D. level scientists that can quickly master the necessary science education research and pedagogical content knowledge than to find education researchers who can achieve the necessary mastery of the science discipline.) These science education specialists then work collaboratively with one or more faculty members to transform specific courses. Through this process, they help the faculty member surmount all three hurdles listed above, and a transformed, more effective course is created. In addition, the science education specialist helps develop educational software, carries out more detailed assessment of student learning, and sets up archival systems so that the materials and results that are produced are saved and can be readily reused and improved upon. Limited data suggests that this approach works even better if there is a team of two or more faculty members working together with the education specialist. These science education specialists also greatly facilitate the discussion and identification of learning goals and coordination of educational efforts within departments. By having a small community of such people working in a variety of science departments, their effectiveness is substantially enhanced; much of the relevant expertise, technology, and methods for research on student learning, as well as knowledge on working effectively with faculty, are highly transferable. It is important to distinguish between instructors, whose primary job is to teach, and these science education specialists whose job is to help faculty change, assess, improve, and disseminate their teaching practices.

It has been noted that successful innovations in education have been characterized by the faculty first becoming aware of the opportunities, particularly with regard to improved assessment, followed by the formation of “learning communities” devoted to development and implementation of improved innovative practices. This implies that transformation needs to address both raising the level of faculty awareness and appreciation of education research, and developing extensive faculty collaboration in the process. The science education specialist provides a means to accomplish these goals with a minimum of energy and initiative on the part of the individual faculty member, and hence reduces the barrier to change.

The design goal of this model is that these science education specialists, and the corresponding increase in resources to support them, will only be required for a limited time. After 5-10 years, sufficient transformation will have taken place and there will be enough gains in effectiveness and efficiency as described above, that the optimization will be self-sustaining at the current level of resources (both financial and faculty time). However, that hypothesis is as yet untested.

VI. Summary.
There are currently great needs and great opportunities for improvement in post-secondary science education. We need to provide a large fraction of our students with complex understanding and problem solving skills in technical subjects. Emerging research indicates that our colleges and universities are not achieving this. However, there are great opportunities to improve this situation using advances in the understanding of how people learn science and advances in educational technology. Realizing these opportunities will require significant changes in how universities approach science education. Change will not be accomplished, however, without investments in assessment of educational outcomes, creation and testing of new educational methods and materials, and the development of faculty expertise in effective education. Strong visionary leadership at the institutional and departmental level will also be required. With adequate investment and leadership, there is reason to hope that we can provide a large fraction of our student population with a far better and more useful education than they are receiving today.

As this is a “think piece” intended to stimulate thought and discussion, rather than a scholarly manuscript, I have not provided extensive references. The references provided are intended in most cases to be representative examples of a more extensive literature. Reference 2 provides a good review of research on science education, and reference 4 has a fairly extensive review of the literature in physics education research.

[1] J. Duederstadt, A University for the 21st Century, Univ. of Mich. Press (2000) provides an extensive discussion of these topics.
[2] J. Bransford et al, How people learn, NAS Press, Wash. DC. (2002)
[3] P. Ross, The expert mind, Scientific American, pg. 64, Aug. 2006, and K. A. Ericsson, et al, The Cambridge Handbook of Expertise and Expert Performance, Cambridge Univ. Press (2006)
[4] E. Redish, Teaching Physics with the Physics Suite, Wiley (2003)
5 W. K. Adams, K. K. Perkins, N. Podolefsky, M. Dubson, N. D. Finkelstein and C. E. Wieman, A new instrument for measuring student beliefs about physics and learning physics: the Colorado Learning Attitudes about Science Survey, Physical Review Special Topics: Phys. Educ. Res. 2, 010101, 2006, and K. K. Perkins, W. K. Adams, N. D. Finkelstein, S. J. Pollock, and C. E. Wieman, Correlating Student Beliefs With Student Learning Using The Colorado Learning Attitudes about Science Survey, PERC Proceedings 2004.
[6] Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Res.*, *15* (2), 4-14.
[7] J. Pellegrino, et al, Knowing what students know; the science and design of educational assessment, NAS Press, Wash. D.C., 2001
[8] see ref. 4 for a review of the content assessment instruments such as FCI, BEMA etc, and ref. 5 for the CLASS.
[9] J. Minstrell, http://www.diagnoser.com.
[10] Koedinger, K., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8,* 30-43.
[11] http://nces.ed.gov/timss/
[12] K. Heller, University of Minn. Physics Dept., and G. Gladding, Univ. of Illinois Champaign-Urbana Physics Dept. Their TA training programs and materials are available on website and/or by request.
[13] Singer et al. (including C. Wieman), Americas Lab Report, NAS Press, Wash. D. C. 2006
[14] M. Fullan, *The New Meaning of Educational Change* (3rd ed. Teach. Coll. Press, N. Y. 2001), and
E. Seymour Tracking the Process of Change in U.S. Undergraduate Education in Science, Mathematics, Engineering, and Technology, *Science Education*, 86 (2002), pp. 79-105.
[15] E. M. Rogers, *Diffusion of Innovations* (5th ed., Free Press, New York, 2003).