ONE

# The Vision

THIS IS A UNIQUE TIME in the history of science education. In recent years, those of us who have called for improvement in Science, Technology, Engineering, and Mathematics (STEM) education are receiving major attention, and there is an increasing awareness that we need to change our approach to the way we teach science. Many of these efforts, including my own, are guided by the work emerging from two rapidly growing fields: the learning sciences and, in particular, discipline-based education research (DBER) in science at the undergraduate level. DBER has produced an extensive body of research with compelling evidence that many of our current ways of teaching undergraduate science, particularly the pervasive lecture, are quite ineffective. Moreover, this new research is laying the foundation for a new model of science education by empirically testing which methods of instruction produce the best results for students. Collectively, these studies indicate that we could significantly improve the quality of science education if universities and colleges adopted these research-based methods on a large scale. Most importantly, these are changes that can be implemented today, and can be implemented within our present institutional structures and, crucially, within our current budgets. I believe, and the Science Education Initiative has gone a long ways toward demonstrating, that by adopting these new approaches to teaching, we can create a higher education system that has most of the same organizational structures and priorities—and the same price tag—as the one that currently exists, but provides far greater educational value. In this chapter, I lay out how such an improved system of higher STEM education would look.

## The Educational Goal

By engaging in study within a discipline, the student should develop expertise in the subject, including problem-solving approaches and skills, habits of the mind, content knowledge, and beliefs about the nature and relevance of the subject. These learning gains should be visible both at the level of an individual course and across a curriculum or program of study as a whole. At the course level, it is important that students move toward such expert-like views of STEM—even if they are non-STEM majors taking a single course to fulfill requirements. The educational goal should be to have these students understand and think about science more in the way a scientist does, including appreciating the scientific process, relating ideas in STEM to real life, and developing curiosity about the natural world. At the program level, it is important that curricula be purposefully aligned, ensuring that courses build on one another to provide ever-deepening mastery of such core competencies.

In the modern world, there is a growing need for technical literacy and skills across the workforce and in public policy decisions.[1] This makes science education important for all students, not just those pursuing careers in science or engineering. A particularly important segment of this population for whom science education is especially important is the fraction who will become the future K-12 teachers.

There is a large and growing body of evidence indicating that post-secondary science education is failing to meet these educational needs. Although there is a particularly large amount of research on how students learn physics and on the shortcomings of conventional instruction, similar results are seen in chemistry and biology.[2] Most students are learning that "science" is a set of facts and procedures 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 science as less interesting and relevant than they did when they started.[3] The typical student is *not* learning to see science the way an expert does, as a set of interconnected, experimentally determined concepts that describe the world. They are also not learning useful conceptbased problem-solving methods that can be applied in novel contexts, as experts do. Below I discuss the reasons for this and how this situation can be changed.

## The Model for Higher Education: Origins and Needed Change

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.[4] The lecture format, which still predominates in STEM teaching today, began before the invention of the printing press, as an efficient way to pass along basic words and information in the absence of written texts. Economies of scale led to this antiquated model expanding to the current situation of a lecturer addressing a group of largely passive students, often several hundred at a time.

Although it is doubtful that this ever was a very effective model for science education, societal changes over the past several decades have shown that it is clearly unsuitable for science education needs today. The most significant of these changes are discussed below.

*Changing needs.* 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, especially in technical fields. These skills are far more important than simple information/knowledge. The employment landscape is also changing rapidly; many current popular jobs are ones that did not exist ten years ago. 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, lacks the individualized challenging exercises and feedback that are critical for acquiring deep understanding and complex problem-solving skills.

*Changing student demographics.* Until a few decades ago, college education was 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 we need a system that can deliver a high-quality education to that large, diverse population. We face an unprecedented educational challenge: the need to effectively teach complex technical knowledge and skills to a large proportion of the total population.

*Changing landscape of higher education.* 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 higher-education faculty. The modern research university now plays a major role in knowledge acquisition and application in science and engineering. Running a research program has become a necessary part of nearly every science and engineering faculty member's activities, and it is 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 labs into commercial products. This brings additional revenue into the institutions and provides highly visible justification for government expenditures on basic research at universities. When they take this step into commercialization, faculty members are often literally running a business, in addition to having the business-management-like responsibilities of operating a university research program. While good arguments can be made for the value of these various faculty activities, the result is a faculty with new sets of demands and responsibilities that largely did not exist in the middle of the last century. These demands, and hence the need to use faculty time most efficiently, must be considered in any discussion of the future of higher education.

*Growing expertise about how people learn science.* While the changes discussed above affect the educational role and environment of the university, there have also been large but less conspicuous changes in our 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.[5] While throughout history there has never been a shortage of strongly held opinions about what "better" educational approaches entail, now there is a solid and growing body of good research supported by extensive data, particularly at the college level in science and engineering, as to which pedagogical approaches work and which do not work. These research-based methods have shown consistent benefits over the traditional lecture in many hundreds of studies across the STEM disciplines.[6] 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 methods and outcomes. An important part of this research is the better delineation of what constitutes expert competence in a technical subject and how this can be more effectively measured.

To briefly summarize a large field: Research has established that people do not develop true understanding of a complex subject such as science by listening passively to explanations. True understanding comes only when students actively construct their own understanding via a process of mentally building on their prior thinking and knowledge through "effortful study."[7] 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.[8] Furthermore, we know that expert competence is made up of several features. In addition to factual knowledge, experts have distinctive 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. Developing these expert competencies, which go beyond the factual, is part of students' path to expertness.

There are important implications of this research for both teaching and assessment. First, the most effective teaching has the student fully mentally engaged with suitably challenging, authentic intellectual tasks that embody all the relevant aspects of thinking to be learned; provides multiple ways of probing their thinking; and offers targeted and timely feedback that guides improvement in their thinking.

Second, meaningful assessment of science learning requires carefully constructed tests that measure the degree to which students have learned to make relevant decisions and solve problems like experts in a given discipline. Test 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 memorized facts and problem-solving recipes that is the (unintended and unrecognized) function of the typical college examination.

*Changes in the state of education-related technology.* The enormous increases in the capabilities of and access to information technology provide 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 information technology currently available is quite limited in both quantity and quality, in part because its design and use are not adequately guided by good pedagogy.

We are now at a watershed in higher education. We are faced with the need for great change, and we have as yet unrealized opportunities for achieving great change. Below I describe the changes and benefits that could be achieved if these opportunities were fully realized.

## An Optimized University

While one might envision an ideal university that has been totally redesigned and has great resources, it is unrealistic to think that such an institution can be created. So instead I will offer a more realizable vision of a much improved university, an *optimized* university. This optimized university will provide the best undergraduate education possible within two basic constraints. The first constraint is that resources in support of higher education will not dramatically increase. The second constraint is that the long-standing structures of disciplines and departments will remain largely intact, as will current broader faculty responsibilities.

The first constraint is simply pragmatic. There is no indication that higher levels of resources are forthcoming for public education. The second has both practical and logical justifications. Where attempts have been made to create universities with dramatically different organizational structures, such as new University of California campuses without discipline-based departments, over time they have effectively reverted to largely traditional structures. I believe there is a basic organizational reason for this. There must necessarily be some organizational unit (that is, a department or some other entity) that oversees the curriculum. This unit must be able to direct the (graduate or undergraduate) career of a student based on its faculty's collective expertise as to what experiences are necessary to support student learning of the content and skills of the field. Thus, while I assume that the labels and orientation of departments will change (as fields continue to evolve because of new directions in science and technology), departments—or some similarly sized organizational entities responsible for education—will and must continue to exist. The need for entities like departments is determined by the limitations of the human brain, as there is a limit to the range of expertise that a diligent person can master. In a typical discipline or department there is a common set of knowledge and expertise that defines it. These elements are continually evolving as new knowledge and corresponding new types of expertise are found to be important for solving certain types of problems. New fields are developed and, necessarily, other aspects of expertise are dropped from the accepted canon, as they come to be seen as less important to the needs of the emerging field. For example, engineering used to be part of the physical sciences, but as engineering techniques and methods became more sophisticated, it was more productive for people focused on engineering-type problems to have a deeper grasp of those methods, at the sacrifice of areas of physics expertise. A group of people with this new set of skills thereby defined a new field of scholarship and subsequently defined what it meant to be properly educated to function well in this field. Of course, engineering itself has since subdivided into more specific fields, as the same basic process has repeated itself. In recent years the range of skills, tools, and knowledge in biology has enormously expanded, with biology departments going through a necessary process of subdividing as more specialization is needed—it is now impossible for an individual to be an expert in all areas of biology, and, correspondingly, no one individual is able to define what students should learn in order to master all areas. Thus some organizational structure like the department, which represents a defined area of expertise that one person can reasonably grasp, will necessarily continue to be the basic educational unit within the university, although the labels attached to these entities will continue to evolve with time.

Table 1.1 outlines characteristics of this optimized university, contrasted with the typical current university.[9]

**Table 1.1.** Differences between current and optimized universities

| Current university | Optimized university |
|---|---|
| Educational focus is on the topics covered and the educational process (for example, number of students taking courses, list of topics, and so forth). Meaningful learning goals are not articulated. | Focus is on the desired student educational outcomes. Learning goals, defining what students will learn to be able to do from courses and programs, are explicitly stated. |
| Instructional model is that the faculty deliver information to the students, who learn it from listening and then practicing in isolation. | Instructional model is based on research on learning. Students must actively practice and develop their capabilities to become more expert, often collaboratively, with ongoing guidance of faculty members. |
| Faculty have sophisticated and extensive content knowledge in their discipline. | Faculty have sophisticated and extensive content knowledge in their discipline. |
| It is assumed that, since the faculty know the content, they know how to teach it effectively. Most faculty are unaware of the relevant research on learning and discipline-based education research. | Faculty have sophisticated pedagogical content knowledge (knowing how the content and skills are best learned, what common student difficulties are and how to overcome those, and how best to motivate students to learn eagerly and effectively). |
| Outcomes are assessed using tests hastily created by individual faculty members and primarily designed to rank students. | Outcomes are assessed by meaningful measures of learning collectively developed by departments, as well as student completion rates per course and program. |

| Academic program design | |
|---|---|
| Each academic program has a series of courses that are required. These requirements are set rather idiosyncratically and revised intermittently when someone in the department feels the inclination. | Each academic program has a clearly delineated set of educational goals that encompass the full set of skills, knowledge, and ways of thinking that are part of an education. These goals are created collectively by the faculty in consultation with other stakeholders, such as industry, educational systems, and government. |
| The department offers a set of courses defined by a list of topics. These choices largely reflect faculty teaching interests and past history. | Each academic program has a series of courses that are carefully aligned and sequenced to progress toward the program goals. Each course is defined by explicit and detailed learning goals that identify the full set of student knowledge and competencies provided by the course. |
| Faculty work in isolation to set their own agendas and goals. | These learning goals explicitly relate to the program goals and are established by a consensus of the department faculty members. They are maintained, regularly reviewed, and updated as part of the normal functioning of academic departments. |
| When students enter into a program, or even just a course, their background preparation is largely unknown and undetermined. Faculty members routinely spend time unnecessarily teaching known material, while also leaving large gaps in coverage of important items. Prerequisites are often poorly matched to what is needed. | The backgrounds of the students in each course will be known and conveyed to the instructor. |

**Table 1.1.** *(continued)*

| Current university | Optimized university |
|---|---|
| Teaching, feedback, and assessment | |
| Faculty use whatever teaching methods they prefer, usually traditional lecture. | In each class, the students encounter pedagogical approaches, materials, and technology based on careful research and measurement of results. Research-based interactive engagement teaching techniques are used extensively, with ongoing feedback provided to students (individually and collectively) during class. |
| Faculty members occasionally ask students which topics they have learned. A few develop review materials for the students based on their best guesses as to what some students lack that is important. | Before a course starts, students complete a detailed diagnostic examination that accurately measures their preparation. This examines their content and conceptual knowledge of the subject and those subjects that the course builds upon, such as mathematics and related science disciplines. It also diagnoses their attitudes and epistemologies about the subject and how it is best learned. |
| Many faculty members spend several class periods rapidly reviewing the knowledge they think the students need. The students who already know the material are bored and find this a waste of time, and often they become overconfident as to the level of challenge of the course. The students who are less prepared don't benefit because the review is too rapid to learn from. | Before a student has ever seen an instructor, the instructor has a profile of his or her strengths and weaknesses, and the computer has already flagged serious deficiencies. If these deficiencies are widespread, the student is guided to enroll in a more appropriate course. |

| | |
|---|---|
| | Where the deficiencies are localized and not severe, the computer provides the student with feedback and suitable exercises to complete that remedy these deficiencies. This ensures that the course will begin with all students at roughly the same level of knowledge and competence, and the instructor has an accurate profile of that level. |
| | The instructor uses the profile of the class to suitably tailor the learning environment. |
| The frequency of evaluation of the students is determined by the instructor and typically includes only graded homework (although often does not), one or two midterm exams, and a final exam. Due to a lack of faculty expertise, many homework problems and faculty-created exams primarily practice and test memorized facts and procedures. Feedback from these evaluations is usually delayed by one to two weeks and provides little to the students beyond a score showing the fraction of questions answered incorrectly. | There are regular ongoing evaluations of the student's thinking and learning throughout the course. These evaluations are linked to targeted timely feedback to both student and instructor. Information technology is used widely to support this ongoing evaluation and feedback, including online homework systems that include intelligent grading and tutoring programs. |
| Typically, technology is developed and used for its own sake, often provides little educational value, and is seldom evaluated as to its effectiveness. | Technology is chosen by looking carefully at how it can enhance learning by supporting good pedagogical design, enhance the capabilities of the instructor, and improve instructional efficiency. |
| Collaborative learning by students is discouraged by curve grading and relies on informal student arrangements. Communication and teamwork skills are usually not part of regular science courses. | The full benefits of collaborative learning are realized by building such collaboration into the structure of the classes, assignments, and grading. This also improves students' teamwork and communication skills. |

**Table 1.1.** *(continued)*

| Current university | Optimized university |
|---|---|
| Faculty teaching evaluations are based on student course evaluations. These have little correlation with learning and none with the use of effective teaching methods. | Faculty teaching evaluations are linked to good measures of student learning and the use of the most effective teaching practices. |
| Saving faculty and student time—improvement in efficiencies | |
| Each new instructor to a course typically reinvents it anew, spending a huge amount of time in the process, and creating a course that is informed neither by past experience at the instructor's own institution nor by experience at other institutions nor by relevant educational research. | Benefiting from experience and saving time by copying what works, course materials and teaching methods are passed from one instructor to the next and continually improved. The relevant discipline-based education research is consulted and good examples of teaching the topics at other institutions are copied. |
| Teaching is an isolated activity in which faculty set their own agendas and goals for the courses they teach. Topic emphasis, teaching methods, and assessments vary widely depending on who is teaching. As a result, there is limited coherence within the curriculum and considerable inefficiency in coverage, wasting both student and faculty time. | Collective ownership is felt within department for courses and curriculum. Collective ownership, along with use of best practices listed above, ensures consistency, coherence, and effectiveness across the curriculum, saving faculty time in preparation and instruction, and making student learning greater and more efficient. |

| | |
|---|---|
| As noted above, the range of student preparation in every course, and the lack of information as to that preparation, results in teaching that wastes large amounts of faculty and student time due to instruction that is redundant for many while being too advanced for others. | Spread and uncertainty in student preparation is reduced. As discussed above, use of diagnostic exams and targeted interventions, and greater educational effectiveness of courses and consistency in learning guided by clear learning goals, will greatly reduce spread in student preparation and provide detailed information to the instructor for every class. |
| Because of uncertainties as to coverage and learning in previous courses and general student preparation, the same science topics are covered repeatedly in the curriculum of a science major, but commonly covered so rapidly each time that the students do not achieve mastery. Such a process is inefficient as well as pedagogically flawed. The research shows that if students develop an incorrect understanding of the material, then shallow repetition tends to reinforce rather than correct such misunderstandings. | Unnecessary repetition is avoided. As discussed above, because of the carefully designed coherence in the curriculum, coverage in each course is optimized to build on what comes before it in the most efficient manner. |
| A large amount of faculty "teaching" time is spent on low-level administrative tasks that could be performed by 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, and so forth. For large classes these low-level tasks can take up a large amount of faculty time. | Expensive faculty time is no longer spent on low-level administrative tasks, particularly in large courses. Dedicated administrative staff, aided by effective software, handle all such tasks more effectively and at lower cost than when done by faculty members. This includes making arrangements for the increasing number of students who have special needs that require adjustments in teaching and/or testing. |

**Table 1.1.** *(continued)*

| Current university | Optimized university |
|---|---|
| Just as faculty receive inadequate training for teaching, so do TAs. TAs are often put in charge of lab and recitation sections and given very little guidance or oversight. A few exceptionally dedicated TAs provide a good educational experience for the students, but the majority do not develop sufficient expertise. | Trained teaching assistants become important contributors to undergraduate education. Well-designed and tested training programs routinely produce extremely well-qualified TAs. Both graduate student and undergraduate student TAs develop a good understanding of how to effectively address student difficulties within the discipline. This provides excellent educational experiences for students and also is the first step in developing science teaching expertise for future faculty and K-12 teachers. This is a low-cost way to improve the student educational experience. |
| Class size is determined by one of three factors: (1) the economics of maximizing the number of students per instructor, (2) a general unsupported belief that there is educational benefit to smaller classes, or (3) the faculty preference for teaching smaller classes, particularly in their areas of specialization, because such classes are more fun to teach, with more personalized interactions with students, and there is less administrative burden with fewer students. This often results in departments creating many small boutique upper-division courses, along with having very large introductory courses. There is no reason to believe that the current class sizes are optimal in any educational or economic sense. Costs per credit hour are usually heavily weighted toward upper-division courses instead of lower-division courses, when good arguments can be made that the opposite would be provide more educational benefit. | Class size is set based on careful measurement and optimization of educational benefits per cost. Technology is used to enhance the most effective research-based interactive teaching in all classes, but it provides the most benefit in larger classes. Technology continues to be developed to help the instructor to make larger classes more intellectually engaging, personalized, and educationally effective. Research is not available to say what class size optimizes learning given a fixed amount of resources and best instructional methods, but the mantra of "smaller is better" is almost certainly not the optimum. There are demonstrations of classes of 200 or more achieving very good learning gains by utilizing technology and research-based practices—with skilled instructors, nearly as good as the best that has been demonstrated in much smaller classes. |

## Issues and Challenges in Optimizing the University

There are some substantial impediments to moving from the current situation to the optimized university. These include structural and administrative limitations, the balance of research and teaching, and failures in the market and in incentive systems.

### Structural and Administrative Limitations

University governing systems are poorly suited to making changes on a time scale that is rapid relative to the faculty life span, which can be several decades. The tendency in the United States toward rather short-lived upper administration (the tenure of public university presidents in the United States now averages less than five years), combined with the pattern of sharing governance with faculty members who have careers lasting decades, effectively puts the administration in a very weak leadership position. In the United States, university governing boards and the position of public university presidents have become highly political and subject to the vagaries of current events, college athletic teams' success, and political intrigue, thereby greatly weakening and distracting administrative academic leadership. Unfortunately, at the same time that administrative leadership is being weakened, modern research universities have grown too much in size and complexity for regular faculty to have all the information and experience needed 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 of desired changes, such as the optimization of undergraduate education described above.

Another closely related complication is that that 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.

## Balance of Research and Teaching

The appropriate balance of teaching and research in the optimized university remains a matter of debate, with no clear best weighting. Both teaching and research are essential components of the modern research university and are vital contributions to society, and to be a highly effective teacher in a discipline, one must be an expert in that field (as well as having expertise in teaching). It would be unwise to abandon either. However, optimizing the use of faculty time offers enormous potential for improvement in educational effectiveness and efficiency. The best approach is to achieve those improvements and examine the results before considering any changes to the current balance of research and teaching. Also, it is hard to imagine that faculty members 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 based only on what they themselves learned in school. Thus maintaining an active research program in a department clearly serves to enhance the desired faculty expertise in teaching.

## Failures in the Market and in Incentive Systems

Teaching in the modern university displays a well-known phenomenon in economics: that free markets do not function properly in the absence of information. In the context of higher education, it is next to impossible for prospective students to get any meaningful information on the quality of teaching at the institutions they are considering. So they are forced to make decisions based on very distant proxies, such as the research productivity of the faculty at a given institution, the cost of tuition, or the quality of the dormitories. Once at the institution, they might be able to make decisions about courses based on student course evaluations, but it has been well established that such evaluations have a host of problems, the most important being that there is no correlation between student evaluations and objective measures of learning,[10] and we have seen no correlation between evaluations and the use of effective instructional practices.[11]

As a result of these information failures, the educational value provided by an institution of higher education, how sought-after it is by prospective students, the amount of public support it receives, and support provided to the faculty who generate that educational value are all completely disconnected. The lack of information results in a lack of incentives to improve educational quality.

The biggest barrier to improving the teaching at research universities is that they are so ineffective at measuring and rewarding effective teaching. There are no incentives for educational change built into the system, and there are several disincentives. Only after the lack of effective measures of teaching quality is addressed will it be possible for prospective students, state governments, the public, and institutions themselves to recognize and reward teaching quality. This will provide the necessary incentives for institutions, and faculty within those institutions, to adopt the best teaching practices and work to improve educational outcomes. We have developed better methods of evaluating teaching as part of the SEI efforts.[12] When measures such as these are in widespread use and the resulting information is available, it will then be possible to have a meaningful incentive system that will drive ongoing improvement in educational quality. This will also allow rational decisions about the appropriate weighting of research and teaching in the optimized university, as well as sensible variations in this weighting across different types of institutions.