What We Know About Learning*

HERBERT A. SIMON
Department of Psychology
Carnegie Mellon University

ABSTRACT

Traditionally, engineering research and teaching have been approached in very different ways. To prepare for research we undergo years of rigorous training, both in scientific knowledge and in methods of gaining new knowledge through experimentation and analysis. To prepare for teaching, we acquire the same knowledge, but, except for a stint as a teaching assistant, we receive almost no training in how to impart it to students.
Fortunately, there is now a well developed science of human learning which has strong implications for the ways in which our students should learn and we should teach. This paper comments on some of the things we know about human learning that can substantially improve our university instruction.

What do we know about learning and the relevance of that knowledge for engineering education? I am pleased to be part of a discussion of this topic. I am also pleased that much of the concern today with instruction in the engineering college focuses on learning about the design process. There was a period not too far back in our engineering schools when we thought we were achieving some kind of prestige in academia by assuring everyone that what we were doing was really engineering science. As a result, it became harder and harder to discover in engineering colleges anyone who was concerned with teaching the process of engineering design.

I think this trend is now reversing itself. Perhaps due in part to the advent of computers and our attempts to automate various aspects of the design process on computers, we have begun to understand that design, like every other human mental activity, has systematic foundations. There is a "science of design," and since it is a science — we can teach it. Of course, if we didn't call it a science we still could teach it; but the fact that we now can analyze the synthesis process — the kinds of mental processes that go on in design — gives us confidence that we can learn how to teach it effectively.

I. LEARNING TO TEACH AND LEARN

Universities are very long-lived institutions and, in several respects, rather curious ones. The University of Paris was founded around the year 1200, when students prepared their own textbooks by copying their professors' lectures. In spite of the invention of printing not too long thereafter, students still continued to behave in their classes as copyists — assiduously taking notes, recording the deathless words of professors as if they didn't know printing had been invented and was available. I have heard that there are some universities where this happens even today.

I would attribute the tardiness in responding to new technology in part to the amateurism of learning and teaching in the universities, then and now. (I would have a different story to tell about K through 12. There may be some problems in the way education is conducted in K through 12, but amateurism is not one of them.) But amateurism is endemic in the university. I mean "amateurism" in a technical sense: to wit, neither we, the faculty members, nor our students have received any significant formal training for doing what we do during much of the day, every day.

Most of us were trained as teachers by serving as TA's in a couple of classes when we were graduate students; and our students, with rare exceptions, have never received any systematic and consistent instruction in how to learn. Yet that is the skill they have been exercising every working day of their lives for more years than they would like to remember. So students don't study the skills of learning, and university teachers don't study the skills of teaching.

I don't mean that learning and teaching doesn't go on in our universities; I think that a great deal does. But there is every reason to believe that one could, by introducing a higher level of professionalism, make both what the students are doing and what we are doing with our students substantially more effective.

Contrast the present practices, just for a moment, with sports, which also go on at universities. Athletes are systematically trained to be athletes, and if you examine the athletic training carefully, you find that much of it is directed toward learning how to acquire skills: how you need to behave as an aspiring athlete so that tomorrow you will be a better athlete. The learning process, however informally it may be handled, is very much a part of the typical athletic training and coaching regime. Coaches themselves are trained in the skills of coaching. Now, maybe we play our games better than they play their games, but I don't think we should count on that. I think we should ask seriously whether we, too, should not be paying explicit attention to the techniques of learning and teaching.

How do we bring that about in a university? I'd like to brag a little about Carnegie Mellon, or as it was called then — Carnegie Tech, because when I arrived, around 1949, there was something in place called the Carnegie Plan. The Carnegie Plan had several aspects: first, it was spawned in the Engineering College, but that wasn't surprising because Carnegie Tech then was mostly an engineering college. The idea of the Carnegie Plan was that the emphasis in engineering education should not be placed on knowledge, but should focus attention on the learning processes and the problem solving processes of the students. The goal of training students

* 1997 Frontiers in Education Conference Keynote Address sponsored by the Lockheed Martin Corporation, November 7, 1997.

was to enable them to execute skills; they had to acquire those skills; and the most essential skills were the broad skills that we call problem solving. You could be sure in an engineering school that you wanted to teach and exercise those skills in the context of engineering problems, but the skills themselves were broader and more basic than that.

Let me tell a brief story about how that came about. Our president at the time was Bob Doherty. Doherty came from General Electric via Yale, and had been one of the bright young men who were taken under the wing of the famous engineer Stiglitz. Every Saturday, Stiglitz would hold a session with these talented young men whom General Electric had recruited and who were trying to learn more advanced engineering theory and problem-solving techniques. Typically, Bob Doherty would sometimes get really stuck while working on a problem. On those occasions, he would walk down the hall, knock on Stiglitz's door, talk to him — and by golly, after a few minutes or maybe a quarter of an hour, the problem would be solved.

One morning Doherty, on his way to Stiglitz's office, said to himself, "Now what do we really talk about? What's the nature of our conversation?" And his next thought was, "Well Stiglitz never says anything; he just asks me questions. And I don't know the answer to the problem or I wouldn't be down there; and yet after fifteen minutes I know the answer." So instead of continuing to Stiglitz's office, he went to the nearest men's room and sat down for a while and asked himself, "What questions would Stiglitz ask me about this?" And lo and behold, after ten minutes he had the answer to the problem and went down to Stiglitz's office and proudly announced that he knew how to solve it.

So you can see that Doherty was very deeply interested in the processes of learning, and it was he who was largely responsible for bringing into Carnegie Mellon the Carnegie Plan, with its emphasis on problem solving and learning to learn. He also did much to bring about the balance between professional and liberal education, which we all take for granted in engineering schools today, but didn't then.

II. Our New Understanding of Teaching and Learning

It's all very well to say that in the university we should pay more systematic attention to the learning and teaching processes; but there is no use doing that unless we have something that can be taught about learning and something that can be taught about teaching. There must an underlying science here, or (if we don't want to use such a prestigious term) some systematic underlying knowledge of what processes have to go on in the human mind when a person is learning and when a person is teaching.

Psychology has been interested in learning processes for a long time, going back to the earliest parts of its history at the turn of the 20th century. At the beginning, taking its cues from advanced sciences, it said, "We have to study the simple before we can study the complex." Therefore, there was a long period of time in American psychology when what you studied were rats — because rats were allegedly simpler than human beings. The idea was that if you really understood how rats learned by running them through mazes, then maybe you would have something to say about how people learned. Of course psychologists did experiments with people as well as rats, and much was discovered about both rat and human learning.

This period of "Behaviorism" is associated with the name of Watson, and later, with that of B. F. Skinner.

More recently, two things have happened to psychology. First, there has been a vastly increased willingness to do research on complex human tasks — not simply memorizing nonsense syllables or simple concepts — but tasks at a professional level. One of the tasks that has been most studied (and it certainly is a professional task) is chess playing. Today we can say a great deal about what a chess player knows and what processes a chess player uses to select good moves. Chess has become sort of the Drosophila for research in cognitive psychology — the standard research organism — the standard research setting in which we have accumulated an enormous amount of knowledge.

There are other even simpler "organisms" we have studied — the puzzle called the Tower of Hanoi, where you move disks from one peg to another to achieve a desired pattern. There again we have found some underlying principles of problem solving activity — something called means-ends analysis (which you're all familiar with even if you are not familiar with the name) that is frequently observed in human problem solving. So, the Tower of Hanoi, being a simpler organism than Drosophila, became the E. coli of research in cognitive psychology.

III. Learning and Teaching Design

Engineering and engineering design have been much affected by these developments. Since the 1960's there has been increasing attention to the design process, and I assume from your presence at this meeting, that many of you are familiar not only with the recent research but with recent textbooks on engineering design (an excellent example is Dym).[1] A parallel process has been going on in architecture, where a lot of good work had been done on the very ill-defined tasks of designing a house or an office. Design has already been a beneficiary of this new trend in psychological research. Another thing that has happened is that increasing numbers of psychologists have become interested in the application to education of what we have learned about learning.

For example, at Carnegie Mellon my colleague, John Anderson, has, for about a decade, been building computer tutoring systems for high school students, primarily in geometry and algebra, and in logic and computer programming. Computer tutors are not a new thing, but what is new, and growing very rapidly, is the design of those tutoring programs on the basis of what we know about the learning process — on the basis of underlying psychological principles.[2]

For example, this means starting the design of the tutor by analyzing the task and the skill that you are trying to teach; if the task is geometry, determining what the student has to have stored in his or her head in order to prove geometric theorems and solve geometry problems. The first thing you must do is to define goals. Then you try to specify the content of those goals in terms of human thinking processes; and then you ask what kinds of experiences, if the students were exposed to them, would lead them to acquire these skills. What is new in this kind of computer tutor is a design based not merely on what computers can do but on what people can do when computers provide them with certain experiences. These are two quite different ways of looking at the technology. As I go along, you will see why I emphasize this point.

When we study the process of design, we discover that design is problem solving. If you have a basic theory of problem solving, then you are well on your way to a theory of design. This discovery came as no deep surprise to those of us who had been doing cognitive research on human thinking in other domains. Design is a special kind of problem solving.[3]

Design is usually the kind of problem solving we call ill-structured. Unlike the Tower of Hanoi or even chess, you don't start off with a well- defined goal. Nor do you start off with a clear set of alternatives, or perhaps any alternatives at all. Goals and alternatives have to emerge through the design process itself: one of its first tasks is to clarify and elaborate goals and to begin to generate alternatives.

If you look at any really complex engineering or architectural design you find that the goals are never completely defined until the design is almost finished. At any time in the process of designing you can say, "There has to be enough space there to allow that door to swing open. So I'll have to set that as a new constraint and make sure that condition is satisfied." The very process of design reminds you of new conditions that have to be satisfied.

A characteristic of design that is special to it, besides this gradual emergence of goals, is that the largest task is to generate alternatives. There are lots of theories of decision making, a field that has been heavily cultivated by economists and statisticians. But most theories of decision making start out with a given set of alternatives and then ask how to choose among them. In design, by contrast, most of the time and effort is spent in generating the alternatives, which aren't given at the outset.

Of course generating alternatives and choosing among them aren't isolated from each other. The process of design is a continual cycle of generating alternatives and testing to evaluate them. The idea that we start out with all the alternatives and then choose among them is wholly unrealistic. If you are designing an important bridge, you might consider two or three basic kinds of bridges and choose one, then go to the next level of detail, and so on. Throughout the design process you are always generating two or three alternatives and choosing among them, and then setting the values of specific parameters to fit the application at hand.

We are beginning to learn what kind of a problem solving process design is, and what its underlying principles are. If we want to teach design effectively, then we as teachers ought to know a good deal about its theory and about the empirical evidence that supports the theory. And these matters ought to enter into the construction of our courses and ought to enter into the construction of the instructional technology that we use in those courses. Let me hint at a few other things we have learned about design and about expertise in general — for the engineer is supposed to be an expert in whatever kind of engineering he or she is practicing.

IV. The Nature of Expertise

There has been a great deal of research on what constitutes expertise.[4,5] In addition to chess, medical diagnosis is a standard environment in which much research on expertise has been done. How does the doctor, looking at the patient who comes into the office, decide what is wrong, produce a Latin name for it, and often, produce some advice about treatment? How is that done? In the cases of both medical diagnosis and chess playing we have discovered how it is done, and I don't know why it should be basically different in the case of engineering. As a matter of fact, we have some knowledge about that too.

First, the expert possesses a large indexed memory in the area of expertise. In every field that has been investigated, the expert has a minimum of about 50,000 to 100,000 "chunks" of knowledge. "Chunk" is a technical term in psychology, meaning any unit of knowledge that has become familiarized and has a place in the memory's index. As it has a place in the index, a chunk is anything you can recognize in your field of expertise. English speakers are experts on the English language — we have stored over 100,000 familiar chunks, which are called words. When we see them in a text, we recognize them and retrieve their meanings from memory.

Now what does "index" mean? An index is a set of patterns that enable you to recognize things about which you have knowledge, whenever they appear. It's no use to have knowledge unless you can get access to it when it's relevant; and getting access to knowledge when it's relevant uses the process we call recognition. If you say "Hi, mom" and someone says, "How did you do that?" you reply, "Well I can recognize my own mom!" We're not very good at telling what features we saw in order to recognize her but we can be sure of the act of recognition.

It has been shown by studies of numerous fields of expertise that a large part of the skill, of say, a doctor when you walk into the office is simply the skill of recognizing patterns. Sometimes we use a fancier word: we say that we do it by intuition. Intuition is essentially synonymous with recognition. Having an intuition means you get knowledge about something without quite knowing how you did it: without knowing the underlying process. Usually, intuitions come rather suddenly, and somebody says "How did you know that?" and you say "Oh, I had an intuition." You would sound a lot less mysterious if you just said, "Oh I recognized it. I recognized that chunk." And having recognized the chunk you do just what you do with the index of an encyclopedia: you get access to all those things you know about it, stored in your brain. That's one large part of what expertise is all about.

Another part of expertise is the skill of searching through a problem space: of searching from the situation you are in now toward a goal situation, and having skills of asking what to do next. This is where means-ends analysis is used. "I am here; I want to be there. What is the difference between here and there? What operators do I have that sometimes reduce differences of that kind? Now lets apply such an operator and see whether we can make progress."

All sorts of artificial intelligence systems have been devised that do intelligent things: they are usually called expert systems. In general, their expertise is much narrower than human expertise; but in many, many cases — increasing numbers every day — we produce computer programs that perform at expert human levels. When you look inside those programs, what you primarily see are (1) ability to recognize familiar patterns, (2) access to information about the implications of those patterns whenever they are recognized, and (3) ability to do a modest amount of very selective search toward the goal.

Of course, with computers we don't have to be as selective in our searches as human beings must be, because computers are much faster in search than human beings. If you examined a fancy artificial intelligence program like *Deep Blue*, the program that beat the world's chess champion, Kasparov, you would find that it is just a combination of 1) a large amount of chess knowledge, indexed in such a way that the computer can recognize important features on the board and draw on its knowledge for moves to deal with such features, and 2) a very large capacity to search ahead.

But don't be misled by that process. The immensely rapid look-ahead process is not itself the thing that makes *Deep Blue* a world champion — in spite of IBM's claims to the contrary. No amount of search would do that without sophisticated chess knowledge to direct the search along promising paths (knowledge about alternative designs) and chess knowledge to evaluate the positions reached (choice evaluation knowledge). That knowledge was provided to *Deep Blue* by chess experts, including a grandmaster consultant. A later generation of computer chess programs will acquire the knowledge directly from chess books and teachers, by learning processes and not by being programmed. That has already been accomplished for other games (e.g., checkers, as early as 1958), but not yet for chess.

V. Some Principles of Learning and Teaching

We knew a good deal about learning long before this kind of research in artificial intelligence and cognitive science was carried out. Some of what we knew was imbedded already in the wisdom of what I earlier called the Carnegie Plan. First principle: learning has to occur in the students. You can do anything you like in the classroom or elsewhere — you can stand on your head — and it doesn't make a whit of difference unless it causes a change in behavior of your students.

Learning takes place in the minds of students and nowhere else, and the effectiveness of teachers lies in what they can induce students to do. The beginning of the design of any educational procedure is dreaming up experiences for students: things that we want students to do because these are the activities that will help them to learn this kind of information and skill. And then we can back off and ask what we have to do to get students to carry out these activities.

VI. The Role of Educational Technology: Computer Displays

Notice that if you proceed in this way, technology is the tool but it is not the driver. What we must avoid above all is designing technologically sophisticated hammers and then wandering around to find nails that we can hit with them. That is a great temptation for all of us who are involved with computer technology; for computers can do really fascinating things when they are not being stubborn; and we would like to see how we can use those potentialities in education. But I submit that we are not going to succeed in that unless we really turn the problem the other way around and first specify the kinds of things students ought to be doing: what are the cost-effective and time-effective ways by which students can proceed to learn. We need to carry out the analysis that is required to understand what they have to do — what activities will produce the learning — and then ask ourselves how the technology can help us do that.

Let me give a crude example of this from real life. We all know that we don't do all of our reasoning in words, we do a lot of it in pictures. One of the uses of computers is to present pictures and displays. When one is going to give a talk on visual reasoning, the first things one is tempted to think about are visual displays to use as examples. But there is an alternative way to go about giving such a talk.

It's fun to give a talk on visual reasoning without using the screen, because each of us has a screen, it's called the "mind's eye" and each of us can visualize things in it. So if we are presenting material about visualization, we can give our students exercises in visualizing. "Here is a rectangle, and here is a diagonal line from its upper-left-hand corner to its lower-right-hand corner." I don't have to draw it at all, because most of you, while I was describing it, already have "drawn" a rectangle in your mind's eye and "drawn" the diagonal crossing it.[6]

If we understand the human mind, we begin to understand what we can do with educational technology. Let me give a negative instance.[6] A couple of years ago, we were interested in knowing how students use visual material in thinking: how a student might come to understand special relativity, in particular, the derivation of the Lorentz equations. We built a high-tech computer display which showed a rod that could be made to move and a light beam that radiated out to the far end of the rod and then reflected back to the near end, following Einstein's original description of the phenomenon in his 1905 paper. We had clocks that were synchronized both with a stationary frame of reference and a moving frame of reference; and we could show the rod in either frame of reference; and the clocks could do their thing; and the students could watch all of this — and it threw them into utter confusion.

Now perhaps we designed the wrong display; in fact, it was the wrong display. Perhaps there was another display that would have done the job. But thinking of it in terms of, "Can we get all of these phenomena on the display? Can we make a virtual reality of it?" was not the right approach. More technology was not what was called for, but more understanding of how people process visual displays.

So then we went back and studied the students a bit: we simply gave them the text of Einstein's 1905 paper to read. If you have ever looked at that paper, you know there's nothing more than algebra on the first seven pages, and yet you get the basic equation out of them that leads to the Lorentz transformation. So we gave students those pages and asked them to read them, and we watched what they did. They began to use their minds' eyes and they drew, on the paper in front of them, diagrams of what they saw in their minds' eyes.

But their diagrams didn't look at all like our computer display. They drew a rod on the paper, and then perhaps they put a little arrow on the end of the rod to indicate that it was moving. After thinking for a few moments, they drew a second rod, displaced to the right. Then they said, "This is where the rod is when the light gets to the far end of it and is reflected." Then they usually drew another arrow showing a ray of light going from its original position when the rod started to move, to the point the far end had reached when the light hit the mirror and was reflected.

If you draw that diagram, I think you will see how, by labeling the various parts you have drawn, you will immediately be able to write the equation that Einstein wrote for the time of the event. Label the arrow for the light (ct) because, at velocity c, that is the distance it went in time t. Label the length of the rod "l"; label the little arrow from the initial position of the rod to where it moved later as the distance the rod had moved, which, with a velocity of v, is vt. Now you see in the diagram that the length of the line ct is equal to the rod length, l, plus the length of the line vt; so you write "$ct = l+vt$" and solve for t.

Thus, the trick with computer displays in teaching is to find out what things human beings are capable of envisioning, how they envision them, and on the basis of that, to design your computer program. The best display for this problem is probably the one the students drew, which was a before-and-after drawing of the critical event (the movement of the light), allowing reasoning about the change. Movement was captured by a stationary display of what economists call "comparative statics."

VII. Learning from Examples

We have found that one of the powerful ways for learning is to be given worked-out examples, step-by-step examples of problem solutions.[7,8] Let the student, by working on these examples, find out how to get from one step to the next. Now, that's a little like learning from doing — throw the student a problem and let him or her solve it. But by providing examples you are allowing the student to solve a series of sub-problems, step-by-step. You can make each step as long or as short as you like, depending on how hard they are for the students.

What does the student do in examining a worked-out example? He or she asks, "What change took place from step one to step two?" Let's say this is an algebra equation to be solved. "Oh, 7 was subtracted from both sides." What difference does that make? "Well, originally we had a number on the left hand side and we didn't want that, so we got rid of it." So the student begins to associate the actions that can be taken with the differences these actions make, which determines the conditions under which one wants to use a particular action.

You have all sat in the classroom of the professor who starts at the left end of the board at the beginning of class, writing equations, and at the end of the class writes QED at the right end of the board. You have been watching carefully, and you know that every step was correct. But what you don't know is why just those steps were selected. Today we know that to understand the reasons for taking those steps, you have to have information about the subject organized in your mind as a set of productions — if-then statements: IF I have a number on the left hand side of the equation and I only want X on that side, THEN subtract that number from both sides.

In computer science we call these IF-THEN statements productions. A human being who is skilled in the domain of a problem operates like a production system, finding what action is appropriate in a given situation (the IF), then applying that action (the THEN). That leads us to ask what kinds of exercises we can give students that will lead them to acquire those productions. When we understand that, we can design a computer tutor, or for that matter a paper and pencil booklet, that will deliver appropriate problems to the students and provide them with these kinds of learning experiences.

As a matter of fact this particular idea has been applied for the last ten years to the teaching of algebra in Mainland China. In several hundred Chinese middle schools, courses in algebra are now taught with almost no lectures, usually in classes of about 50 students. The teacher works in the class, tutoring individual students, and the skills are taught almost entirely by having the students work out examples. Assessments have shown this method to be very effective.

My comments about diagrams and about learning from examples illustrate the interactions between understanding human mental processes, and understanding how to use the technological goodies that are now available to us. Our knowledge of the thought processes has to be brought together with our knowledge of the technology if we are to design educationally effective activities and experiences for our students.

VIII. Applications in the University

For more than 20 years, we have had a Teaching Center at Carnegie Mellon to expose our faculty to principle-based teaching methods of the kinds I have been describing. For about five years, we have had a Center for Innovation in Learning which is applying these kinds of principles within the university by bringing together faculty possessing the important cognitive and psychological skills with the faculty in the departments who are responsible for planning and teaching the courses. In both of these related activities, our aim is to begin to professionalize the educational process in our university. We are convinced that we can no longer afford university education that is provided by amateur teachers to amateur students. In the two organizations we have created, modern educational technology is playing an important and increasing role, but always in the context of sound psychological analysis of the learning that has to take place. We think this is a model that is generally useful.

IX. Conclusion

Modern information technology, including the technologies of computers and communication networks, has great importance for engineering education. First, it has enabled us to create a significant theory of design that provides a structure for teaching engineering design as well as engineering science.

Second, information technology, applied to computer simulation of human thought processes, has greatly advanced our understanding of the nature of expertise and of the learning processes through which professional competence and expertise are acquired.

Third, information technology, and especially its psychological component, enables us to construct new instructional procedures by careful analysis of the structures of the tasks to be mastered, the systems of productions (if-then rules) that underlie successful performance, and the kinds of activities that will enable students to acquire these productions.

Finally, we can find in our computing technology new capabilities — for example, capabilities for visual displays — for presenting to our students the learning experiences that our cognitive theory indicates will be effective for instruction.

We now have the basic capabilities, if we will use them, for transforming university education from an activity for gifted amateurs, teachers and students, to an activity based on solid scientific knowledge, that can be practiced in a thoroughly professional, and hence increasingly effective way.

References

1. Dym, C. L., *Engineering Design: A Synthesis of Views*, Cambridge University Press, New York 1994.

2. Anderson, J. R., et al., "Cognitive Tutors: Lessons Learned," *The Journal of Learning Sciences*, vol. 4, 1995, pp. 167-207.

3. Simon, H. A., *The Sciences of the Artificial* (3rd. ed.), The MIT Press, Cambridge, MA 1996, especially chapters 5 and 6.

4. Ericsson, K. A., and J. Smith (eds.), *Toward a General Theory of Expertise*, Cambridge University Press, New York 1991.

5. Ericsson, K. A. (ed.), *The Road to Excellence: The Acquisition of Expert Performance*, Erlbaum, Mahwah, NJ 1996.

6. Tabachneck-Schijf, H. J. M., and H. A. Simon, "Alternative Representations of Instructional Material," in D. Peterson (ed.), *Forms of Representation* (pp. 28-46), Intellect Books, Exeter, UK 1996.

7. Zhu, X., and H. A. Simon, "Learning Mathematics from Examples and by Doing," *Cognition and Instruction*, vol. 4, 1988, pp. 137-166.

8. Zhu, X., et al., "Cue Recognition and Cue Elaboration in Learning from Examples," *Proceedings of the National Academy of Sciences, USA*, vol. 93, 1996, pp. 1346-1351.