Millikan Lecture 1994: Understanding and teaching important scientific thought processes

Frederick Reif
Center for Innovation in Learning, and Departments of Physics & Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

I. INTRODUCTION

Physics is an intellectually demanding discipline and many students have difficulties learning to deal with it. Further, our instruction is often far less effective than we realize. Indeed, recent investigations have revealed that many students, even when getting good grades, emerge from their basic physics courses with significant scientific misconceptions, with prescientific notions, with poor problem-solving skills, and with an inability to apply what they ostensibly learned.[1–4] In short, students' acquired physics knowledge is often largely nominal rather than functional.

This situation leads one to ask: Why is this so, and what might be done about it? More specifically, it has led me to address the following two basic questions: (a) Can one understand better the underlying thought processes required to deal with a science like physics? (b) How can such an understanding be used to design more effective instruction?

These are the questions which have been the focus of my work during the last several years and which I want to discuss in the following pages.

A. Formulation of the instructional problem

Instruction is a problem that requires one to transform a system S (called the student) from an initial state S_i to a desired final state S_f where S can do things which S could not do initially. This transformation process can schematically be expressed in the form

$$S_i \rightarrow S_f. \tag{1}$$

Although this may seem like a cold-blooded physicist's way of formulating the instructional problem, it is certainly *not* dehumanizing. On the contrary! Rather than dealing primarily with physics subject matter or curriculum, it focuses central attention on the human student S trying to deal with physics.

More important, the formulation (1) of the instructional problem makes apparent that a systematic approach to instruction needs to address the following issues.

(1) Analysis of desired performance (S_f). (a) One needs to specify clearly the desired final student abilities and observable performance. (b) On a more theoretical level, one needs to understand the underlying cognitive mechanisms (knowledge and thought processes) required to achieve these abilities.

(2) Analysis of the initial student (S_i). (a) One needs to describe adequately the characteristics and performance of students coming to instruction. (b) On a more theoretical level, one needs to identify what they know and how they think.

(3) Useful comparisons. A good analysis of desired performance (i.e., of S_f) allows several useful comparisons: (a) A comparison with actual expert performers. (This can suggest improved models of good performance, can help reveal "tacit knowledge" of which experts are unaware, and may sometimes disclose that experts are far from perfect.) (b) A comparison with novice students. (This can reveal anticipated learning difficulties and identify more precisely what needs to be taught.) (c) A comparison with prevailing methods of instruction. (This can reveal the deficiencies of such instruction, e.g., important skills that are never explicitly taught.)

(4) Design of instruction (the transformation process $\rightarrow$). (a) One needs to design an effective learning process whereby the student can acquire the knowledge and thinking skills required to achieve the desired final performance. (b) Finally, one needs to implement this design in practical settings.

The preceding approach to instruction is centrally based on an adequate understanding of the thought processes leading to the desired performance. The basic premise is that one cannot teach physics effectively without an adequate understanding of the thought processes needed in this field (no more than one can teach someone how to play good chess without an adequate understanding of the thought processes needed to play that game).

B. Outline of important issues

Let me then follow the preceding instructional approach to identify some of the specific issues important to the teaching of physics.

Instructional goals. The choice of instructional goals is a matter of judgment and depends also on the particular student audience. However, my central goal has been to help students acquire a modest amount of basic knowledge which they can *flexibly use*. There are at least two reasons why such flexible usability seems centrally important. (a) The goal of science is not the accumulation of various facts, but the ability to use a small amount of basic knowledge to predict or explain many diverse phenomena. (b) Students will have to function in a complex and rapidly changing technological world where they will profit little from knowledge that is rotely memorized or poorly understood. Any acquired physics knowledge will be useful to them only if it allows them to cope flexibly with any future courses or tasks encountered by them.

Abilities facilitating flexible usability. What kinds of thought processes are required to ensure that scientific knowledge can be flexibly used? My work suggests that the cognitive abilities summarized in Fig. 1 are of particular importance. These include the basic abilities required to interpret properly scientific concepts and principles, to describe knowledge effectively, and to organize it effectively. These are necessary prerequisites for more general problem-solving abilities, including the abilities to analyze problems, to construct their solutions, and to check these solutions.

Overview of this paper. In the following pages, I shall examine more closely each of the preceding abilities, pointing out why each of these is important and more complex than one might naively believe. In each case, consideration of the instructional problem $S_i \rightarrow S_f$ will lead me to do the following: (a) Indicate some common inadequacies of students' initial abilities. (b) Analyze the thought processes re-

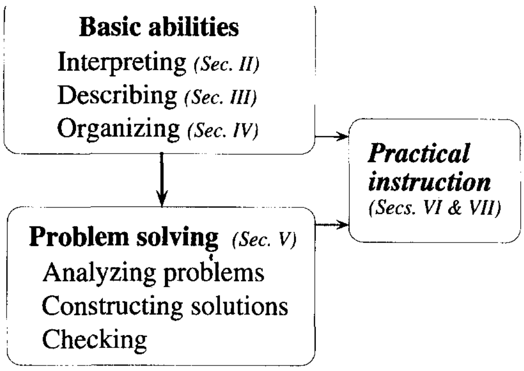

Fig. 1. Central cognitive issues important for scientific work.

quired to achieve the desired abilities to be finally acquired by students. (c) Examine the instructional implications for designing an effective learning process.

As outlined in Fig. 1, Sec. II will examine the interpretation of scientific concepts and principles, Sec. III will deal with effective methods of description needed for scientific work, and Sec. IV will describe useful forms of knowledge organization. Section V will then discuss problem solving (i.e., methods for analyzing problems, for constructing their solutions, and for checking these). The examination of the preceding issues prepares one to consider how all of them may be jointly incorporated in the design of practical instruction. Correspondingly, Sec. VI will describe work done to implement such practical instruction and Sec. VII will mention some of the difficulties faced by attempts at effective implementation. Finally, Sec. VIII will summarize the discussion with some brief concluding remarks.

II. INTERPRETATION OF SCIENTIFIC CONCEPTS OR PRINCIPLES

The basic building blocks of scientific knowledge are special concepts and principles. These are ordinarily quite abstract in order to provide the desired scientific generality (i.e., to ensure that very few such concepts or principles are sufficient to predict and explain many diverse phenomena).

Abstractness as such does not present undue difficulties to people. Many concepts commonly used in everyday life are also quite abstract (e.g., love, truth, beauty, justice, etc.). The difficulty is that one must be able to interpret a scientific concept unambiguously in any particular instance, a requirement *not* imposed on everyday concepts. For example, in daily life there may be many disagreements about whether something is a case of "true love" or whether a particular action is "just." But scientific work does not tolerate similar ambiguities about the proper identification of a scientific concept.

In the sense in which I use it, "interpreting a concept" means identifying or generating the concept in any particular instance. For example, suppose that somebody tells me that a triangle is a three-sided polygon. However, the person cannot recognize a triangle among some other geometric figures, nor construct a triangle with three sticks. Then I would say that the person has some nominal knowledge about a "triangle," but does *not* know how to interpret this concept.

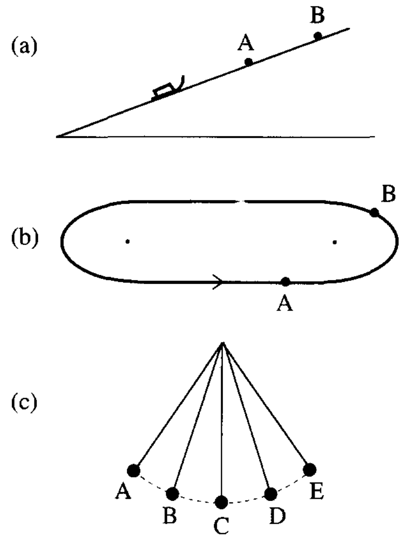

Fig. 2. Situations used for testing the interpretation of acceleration. (a) Sled sliding along a hill. (b) Car traveling around a horizontal racetrack. (c) Oscillating pendulum bob.

The ability to interpret a scientific concept is clearly an essential prerequisite for using the concept to make complex inferences or to do any scientific work with it. Hence one may ask the following question: How well can students interpret the physics concepts which they have ostensibly learned?

A. Observed interpretation deficiencies

To examine this question in some detail, let me focus specific attention on the concept "acceleration." This is a very basic concept, of fundamental importance in Newtonian mechanics and commonly taught at the beginning of any introductory physics course. The concept is specified by its familiar definition that "acceleration is the rate of change of velocity with time," a statement which can also be summarized by the equation

$$\mathbf{a}=\frac{d\mathbf{v}}{dt}\,. \tag{2}$$

Someone able to interpret the concept acceleration should be able to identify the acceleration of a particle in various specific cases, such as those illustrated in Fig. 2. For instance, Fig. 2(a) shows a sled which moves up along a hill, passes the point A with decreasing speed, comes momentarily to rest at the point B, and then slides down again. Figure 2(b) shows a car passing the points A and B while moving with constant speed around a horizontal racetrack. Figure 2(c) shows an oscillating pendulum bob which is momentarily at rest at the extreme point A of its circular arc, passes the point B with increasing speed, reaches its maximum speed at its lowest point C where the string is vertical, continues past the point D, and is again momentarily at rest at the point E.

In a study carried out by me and some co-workers, we presented 15 such specific situations to various persons and observed their responses in detail. The person was asked to

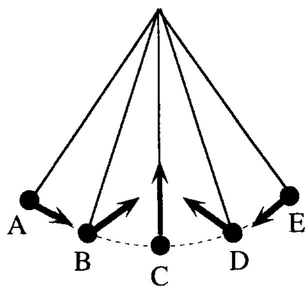

Fig. 3. Accelerations of a pendulum bob.

specify whether the acceleration is zero at the indicated points, or to specify its direction if it is nonzero.

The observed individuals were either students or professors at the University of California at Berkeley. The students, enrolled in an introductory college physics course for prospective scientists or engineers, had been working with acceleration for at least two months. The professors had all taught an introductory physics course in the recent past.

The main results of this study (discussed at length in a paper by Reif and Allen)[5] were the following: The students could answer correctly at most only 35% of such questions. The professors were very much better, but not perfect. (For example, one of them answered correctly only 10 of the 15 questions.)

Recent observations at the University of Washington[6] support these conclusions by more extensive data about various individuals presented with the pendulum problem of Fig. 2(c). Of 124 students who had studied acceleration in the introductory physics course, none could answer this problem correctly; of 22 graduate-student teaching assistants, only 15% could answer it correctly; and of 11 graduate students on their Ph.D. qualifying examination, only 20% could answer it correctly. (Even some experienced physicists have difficulty identifying the acceleration of the pendulum bob. The arrows in Fig. 3 indicate the directions of these accelerations.)

The preceding data indicate that the proper interpretation of a scientific concept is no easy task and that many students do not acquire the ability to interpret the scientific concepts supposedly learned by them.

What are some of the reasons for the observed interpretation deficiencies? One common reason is that students retrieve remembered or plausible knowledge fragments which are often incorrect and which are rarely checked against a definition of the concept. For example, many students deem it obvious that a particle's acceleration is zero whenever its velocity is zero. Or they simply retrieve the fact that the acceleration in circular motion is directed toward the center (without heeding the fact that this is only true if the speed is constant).

Even when students do invoke the definition of a concept, they are often unable to interpret it properly. For example, one student, when considering the acceleration of the pendulum at the extreme point A of its swing, said the following:

"The velocity is zero, so the acceleration has to be zero. Because acceleration equals the change in velocity over the change in time···I mean, acceleration is the derivative of the velocity over time. And the derivative of velocity is zero."

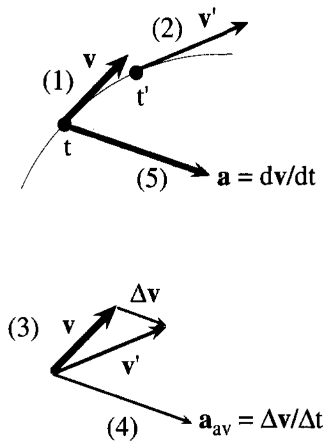

Fig. 4. Defining method for acceleration.

Thus the student invoked explicitly the definition of acceleration, even phrased in the formal mathematical language of a derivative, yet misinterpreted it to reach the wrong conclusion.

B. Cognitive analysis

The preceding detailed studies dealt with acceleration, but yield results consistent with students' observed misinterpretations of many other physics concepts.[1–4] To understand better the difficulties of concept interpretation and the reasons for misinterpretations, let us examine more closely the thought processes required for the proficient interpretation of scientific concepts.

Reliably accurate interpretation. The unambiguous specification of a scientific concept requires that the concept be explicitly specified with sufficient precision so that it can be properly interpreted in any specific instance. This requires an interpretation *method* (i.e., *procedural* knowledge) which specifies what one must actually *do* to identify or construct the concept in any particular instance. (The unambiguous specification of a scientific *principle* requires similarly an operational interpretation method.) The deliberate application of this method can then ensure the reliably accurate interpretation of the concept.

For example, the acceleration of a particle is specified by the defining statement (2). But this definition of the concept is inadequate without the specification of a corresponding interpretation method which involves the following five main steps (illustrated in Fig. 4).

(1) *Original velocity* $\mathbf{v}$. Identify the velocity of the particle at the time t of interest.
(2) *New velocity* $\mathbf{v}'$. Identify the velocity of the particle at a slightly later time t'.
(3) *Change of velocity* $\Delta\mathbf{v}$. Find the velocity change $\Delta\mathbf{v}=\mathbf{v}'-\mathbf{v}$ of the particle during the small time interval $\Delta t=t'-t$.
(4) *Average acceleration* $\mathbf{a}_{av}$. Find the ratio $\Delta\mathbf{v}/\Delta t$, the "average acceleration" of the particle during the time Δt.
(5) *Acceleration* $\mathbf{a}$. Determine the limiting value approached by the average acceleration if the time t' is chosen very close to t (so that Δt becomes infinitesimally small and

can be denoted by dt). The resultant ratio $d\mathbf{v}/dt$ is then called "the acceleration of the particle at the time t."

Application of this method is sufficient to determine the acceleration in all the situations (like those in Fig. 2) presented to the Berkeley students. Note how complex this method really is (especially the third step involving a vectorial subtraction of velocities and the fifth step involving a limiting process)! Yet all this complexity is hidden by the seemingly simple Eq. (2) or by the equivalent statement that "acceleration is the rate of change of velocity with time." No wonder that students have so much difficulty interpreting what this statement really means!

Efficient interpretation. The preceding interpretation knowledge is "formal," i.e., general, precise, and explicitly specified by a method for interpreting the concept in any particular instance. Deliberate application of this formal knowledge can ensure reliable accuracy, but it can be quite laborious. Thus one would also like to be *efficient*, i.e., able to interpret a concept rapidly and with little mental effort.

Cognitive efficiency is not just a luxury for people who prefer to be lazy and save time. It can also be essential for *effective* performance. Indeed, if we always had to spend much time and effort interpreting every concept, we would not have enough mental capacity left to deal with the more complex aspects of tasks in which these concepts are used. (Similarly, suppose that we had never learned to decode individual words and phrases more efficiently than six-year olds. How then would we have enough mental capacity left to read and understand an article in the *Physical Review?*)

Efficient concept interpretation can be achieved by compiling a repertoire of knowledge about special cases of the concept. An encountered situation which matches such a special case can then be recognized almost automatically. Hence such "compiled knowledge" can often be used to interpret the concept intuitively without the need for deliberate processing.

For example, most physicists have compiled knowledge about the acceleration of a particle in some special cases, such as that of circular motion with constant speed. When encountering a particle moving in this way, they then immediately recognize this situation and conclude that the particle's acceleration is directed toward the circle's center. All this is quickly done *without* any need to invoke the definition of acceleration or to engage in reasoning based on it.

Complementary use of formal and compiled knowledge. To interpret a concept both accurately and efficiently, general formal knowledge and case-specific compiled knowledge are used in complementary fashion. If one encounters a familiar situation which matches a particular case in one's compiled knowledge, then this compiled knowledge can be immediately applied. But if one encounters an unfamiliar situation, or runs into puzzling difficulties, or needs to make general arguments, then it is best to invoke one's formal knowledge and to reason from it.[7]

C. Instructional implications

Instruction must ensure that students can adequately interpret any concept or principle *before* they are asked to use it to perform more demanding problem-solving tasks. Otherwise, students are forced to deal simultaneously with the difficulties of concept interpretation as well as with other complexities, a situation which can transcend their learning capabilities and lead to frequent mistakes.

Explicit teaching of interpretation methods. The preceding analysis of the interpretation process suggests the following instructional strategy for teaching a scientific concept (or principle). (a) After motivating and introducing the concept, specify it explicitly together with the associated method required for its interpretation. (b) Let students themselves apply this method consistently in various special cases, including cases which are error prone. (Such error-prone cases include those which require fine discriminations, or which invite confusions with prior notions familiar from everyday life or earlier schooling.) (c) Ask students to summarize the results of their concept interpretations in these special cases so that they acquire a useful repertoire of compiled knowledge about the concept.

There is evidence that the preceding instructional method can be quite effective. For example, by applying this method in an experimental situation, we succeeded in substantially improving students' ability to interpret properly the concept acceleration (from about 40% correct interpretations before instruction to over 90% afterwards).[8,9] The method has also proved quite effective in actual classroom situations.

D. Assuring reliable compiled knowledge

As already mentioned, it is very useful to have compiled scientific knowledge about various specific situations so that these can be quickly recognized. In this way one can develop good scientific intuitions, and does not always need to engage in laborious reasoning from basic definitions or principles.

Need for quality assurance. Such compiled case-specific knowledge can, however, be unreliable unless it satisfies the following conditions: (a) It must be consistent with formal scientific knowledge. (b) It must be carefully discriminated from other kinds of intuitive knowledge used in everyday life or other contexts. (For example, the concept acceleration in physics has properties quite different from those associated with the word acceleration used in conversations with a taxicab driver.)

Most important, it is essential that one be able and willing to check whether intuitively applied compiled knowledge has, in fact, been correctly applied (as judged by consistency with explicitly specified definitions or principles). Otherwise, it is all too easy to ignore fine discriminations or validity conditions restricting the applicability of case-specific knowledge.

Instruction needs to ensure that the preceding conditions are achieved. This is no easy task, especially since students come to science from everyday life where intuitively used knowledge is not subjected to equally stringent requirements.

Fallibility of recognition processes. The ability to recognize a familiar or analogous situation can make concept interpretation fast and effortless. However, recognition processes can be error prone since they do not involve an explicit specification of which particular features should be heeded and which ones can be ignored. This is why recognition processes used in science need to be checked against more reliable interpretation methods. (Hence it can also be dangerous to introduce physics concepts by mere examples or analogies, without more explicitly specified definitions.) The following are some examples.

Figure 5(a) shows three vectors, of equal magnitudes, whose sum is zero. When students are shown the equilateral triangle in the first diagram and are asked about the angle between the vectors **A** and **B**, many say that this angle is 60°.

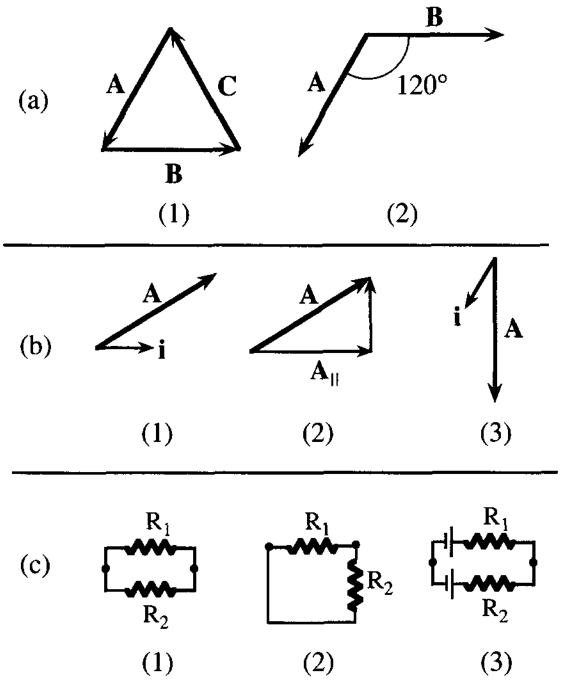

Fig. 5. Fallibility of interpretation by recognition. (a) Angle between two vectors. (b) Component vector of a vector. (c) Resistors connected in parallel.

Of course, it *looks* that way! But this angle is really 120°—if one recalls that the angle between two vectors is defined as the angle between their representing arrows *drawn from the same point* (as indicated in the second diagram).

Figure 5(b) shows a vector **A** and a direction specified by **i**. When students are shown the first diagram and asked to draw the component vector of **A** parallel to **i**, most students can readily do this (as indicated in the second diagram). But when students are asked the same question about these vectors oriented as shown in the third diagram, many have difficulties or answer incorrectly. The reason is that most students try to answer the question by matching against the well-recognized case where the reference direction **i** is horizontal or vertical on the page. The first diagram matches this prototypical situation, but the third diagram does not. Hence the difficulties. (Both situations are almost equally easy if one applies an interpretation method specifying how to construct the component vector of a vector.)

Figure 5(c) shows two resistors R_1 and R_2 connected in three different ways. Many students claim that the resistors are connected in parallel in the first and third diagrams, but not in the second. (Indeed, the first diagram is a prototypical drawing of parallel connection. The third diagram *appears* to match this drawing, but the second diagram does not.) Of course, if one deliberately examines how corresponding terminals are connected, it becomes clear that the resistors in the second diagram are also connected in parallel, but that those in the third diagram are not.

III. KNOWLEDGE DESCRIPTION

Any situation can be described in various ways. For example, it may be described in terms of different concepts, with different symbolic representations (e.g., in words, pictures, or mathematical symbols), and with different degrees of precision. However, such different descriptions are *not* functionally equivalent since the performance of a given task may be facilitated by one description but hindered by another.

Descriptions should, therefore, be chosen so as to facilitate performance of the tasks of interest. This guideline implies that one must know what kinds of descriptions are useful for various kinds of tasks, and that one must know methods for implementing these descriptions.

The following paragraphs discuss how description is important for the interpretation of concepts or principles, and how both quantitative and qualitative descriptions are essential for scientific work. Subsequent sections will indicate that description plays also an important role in problem solving.

A. Description facilitating interpretation

The interpretation of a concept or principle requires that all the ingredients necessary for such interpretation be properly described—with all the concomitant knowledge necessary for such description. For example, in order to interpret the concept acceleration in accordance with the method specified in Sec. II B, one must be able to describe the velocities of a particle at two neighboring times and the difference of these velocities. Correspondingly, one must also have prerequisite knowledge about the properties of the velocity (e.g., that it is tangent to a particle's path) and about the subtraction of vector quantities.

The description knowledge needed for the interpretation of a principle may be even more complex since such a principle relates several concepts, all of which must be coherently described.

Description needed for interpreting Newton's law. To illustrate the preceding remark, consider Newton's mechanics law $m\mathbf{a}=\mathbf{F}_{tot}$. To interpret this law in any particular instance, one must determine the mass m and acceleration **a** of the particle of interest, and then relate these quantities to the total force $\mathbf{F}_{tot}$ (obtained by adding vectorially all the individual forces on the particle). This can only be done if one has first adequately described the mass of the particle, its motion, and the forces due to its interactions with all other objects. Indeed, any deficiency in this description will lead to a misinterpretation of Newton's law—and thus to an incorrect solution of any problem in which this law is applied.

Observed description deficiencies. The preceding description task is far from trivial and often inadequately performed, even by experienced students. For example, in an investigation carried out by Heller and me some years ago,[10] we observed students who had recently completed the introductory mechanics course with grades of B or better. When these students were given some problems of the same kind as those previously encountered in their course, they could successfully solve only a third of them. The reason was that in half of the problems the students faultily described the acceleration or forces, and thus incorrectly implemented the application of Newton's law.

As another example, Shaffer and McDermott[6,11] recently asked various students at the University of Washington to consider a child sitting on a swing at the instant when the child is moving horizontally (passing the lowest point where the rope supporting the seat of the swing is vertical). The students were asked to identify all the forces on the child and on the seat, as well as the total force on each. Of 79 students, who had worked through a special tutorial on forces in an

* **Separate system**
* **Mass**
* **Motion** (velocity & acceleration)
* **All forces**
    * ***Long-range*** (e.g., gravity)
        * Interacting objects?
        * Forces on system
    * ***Contact***
        * Touching objects? (Mark & label contacts)
        * Forces on system
* **Components**

Fig. 6. Method for describing a system by a system diagram.

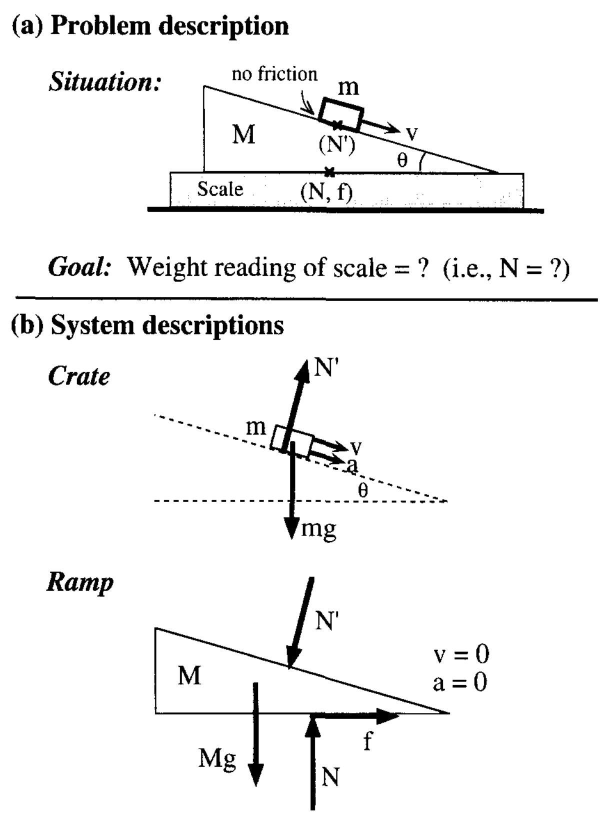

Fig. 7. (a) Useful description of the problem in Table I. (b) System diagrams describing the crate and the ramp (with vectors not yet decomposed into convenient components).

introductory physics course, only 20% correctly answered these questions. Further, of 21 graduate students on their Ph.D. qualifying examination, only 15% correctly answered these questions. For example, more than 50% of these graduate students claimed that the force exerted on the swing seat by the child is just the weight of the child (thus ignoring the effects of acceleration, or distinctions between gravitational and contact forces).

Cognitive analysis. As these data indicate, it is *not* a simple task to describe a system adequately to permit correct application of Newton's law. An attempt to analyze the thought processes required to generate reliably correct descriptions leads to the description method schematically summarized in Fig. 6 and more fully discussed below. This method describes each system (particle or composite particle) graphically by a "system diagram."

To illustrate this method more concretely, it may be useful to consider its application to a specific mechanics problem, such as the one stated in Table I. Figure 7(a) provides a clearer and more pictorial description of this problem. When the description method of Fig. 6 is applied to the crate and to the ramp, these become described by the system diagrams indicated in Fig. 7(b).

The method outlined in Fig. 6 specifies important details of the description process. In particular, it transcends the drawing of conventional "free-body diagrams" in the following significant respects.

(a) Describing both motion and interactions. The method specifies that one describe equally carefully the system's motion (velocity and acceleration) as well as all the forces on the system. This is important because Newton's law, like any other mechanics principle, specifies the relation between motion and interactions. (The omission of motion information from conventional free-body diagrams is thus rather strange.)

Table I. Statement of a mechanics problem.

A crate, of mass m, slides with negligible friction down along a stationary ramp lying on a scale. The ramp has a mass M and its upper surface is inclined at an angle θ from the horizontal. What is the weight reading indicated by the scale?

This provides also a powerful check of the description since one can then verify whether the direction of the acceleration is consistent with that of the total force. [For example, students often forget the friction exerted on the ramp by the scale. But this omission is easily detected in Fig. 7(b) by noting that the acceleration of the ramp could then not be zero.]

(b) Identifying interacting objects before forces. The description method identifies interacting objects *before* specifying the corresponding forces exerted by them. This helps to avoid students' invocation of nonexistent mystical forces (e.g., "centrifugal forces") due to no discernible objects.

(c) Separating long-range and contact interactions. The method clearly separates long-range and contact forces. It also ensures the reliable enumeration of all contact forces by explicitly marking in the situation diagram, like that of Fig. 7(a), all the contact points (and thus corresponding contact interactions).

(d) Labeling contacts by the magnitudes of corresponding forces. By explicitly labeling contact points by the magnitudes of the corresponding forces, one provides a common label for a pair of reciprocal forces. One thus also ensures that these forces, when appearing in different system diagrams, are automatically characterized by the same magnitude.

Such seeming details are, in fact, very important to the generation of correct descriptions leading to proper interpretations of Newton's law. Indeed, the previously mentioned investigation[10] showed the following. When students follow a general description method like that indicated in Fig. 6, their applications of Newton's law (and solutions of problems based on this law) are correct in more than 90% of cases. However, student's performance deteriorates markedly if the description method is less detailed.

B. Complementary quantitative and qualitative descriptions

As already mentioned, descriptions should be chosen to facilitate the tasks of interest. This guideline has some broader implications for the modes of description needed for scientific work.

Needs for precision and extensive inferences. Science aims to explain or predict the maximum number of observable phenomena on the basis of a few basic premises. Especially in a well-developed science like physics, there is then a great need for making extensive inferences.

Formal modes of description, using precisely defined symbols and explicit rules for their manipulation, are very well suited to facilitate long and accurate inference chains. Such formal modes of description, exploiting mathematics and logic, are thus widely used in physics. Familiar examples are algebra, calculus, vector analysis, and many others.

Correspondingly, instruction in physics often places primary emphasis on such formal quantitative modes of description. But are these sufficient for scientific needs?

Need for search. In science, as well as in many other domains, there is a great need for *search*, i.e., for identifying relevant alternatives and deciding among them. For example, search is important for planning approaches to problem solving, for exploring, for designing, for inventing, for discovering, for identifying possible reasons for observed phenomena, for troubleshooting, for exploiting methods of progressive refinement, and for many other such tasks.

Formal methods are *not* particularly well adapted to facilitate such search tasks. Indeed, search is often better accomplished by nonformal methods which use approximate qualitative descriptions expressed in words or pictures. These are methods commonly used in everyday life, and they can also be very useful in science.

Complementary use of quantitative and qualitative descriptions. Scientific effectiveness requires both precise inferences and search. Correspondingly, it can be achieved by the complementary use of both formal quantitative descriptions and nonformal qualitative descriptions. Indeed, good scientists are well aware of this.

For example, Einstein states:[12] "(The physicist's work) demands the highest possible standard of rigorous precision in the description of relations, such as only the use of mathematical language can give." Yet elsewhere, in a letter to Hadamard, he says:[13] "The words of the language ... do not seem to play any role in my mechanisms of thought. The psychical entities which seem to serve as elements of thought are certain signs and more or less clear images which can be "voluntarily" reproduced and combined ... before there is any connection with logical construction in words or other kinds of signs ... The play with the mentioned elements is aimed to be analogous to certain logical connections one is searching for."

As another example, Hans Bethe makes the following comments:[14] "From Fermi I learned ... to look at things qualitatively first and understand the problem physically before putting a lot of formulas on paper. ... Fermi was as much an experimenter as a theorist, and the mathematical solution was for him more a confirmation of his understanding of a problem than the basis of it."

Similarly, Feynman talks about himself in the following vein:[15] "What I am really trying to do is bring birth to clarity, which is really a half-assedly thought-out pictorial semi-vision thing. ... It's all visual. It's hard to explain. ... Ordinarily, I try to get the pictures clearer, but in the end the mathematics can take over and be more efficient in communicating the idea of the picture. ... In certain particular problems that I have done it was necessary to continue the development of the picture as the method before the mathematics could be really done."

C. Instructional implications

The preceding comments imply that instruction must deliberately foster students' abilities to describe their acquired knowledge in useful ways.

Teaching description methods. When students are expected to perform important tasks (like interpreting scientific principles), one needs to identify in sufficient detail the descriptions facilitating these tasks and to teach explicit methods for generating such descriptions.

Indeed, skills of description and interpretation are sufficiently complex that they deserve to be taught in their own right *before* students are asked to use them in more demanding problems. For example, it is profitable to spend time when students can merely learn how to describe the motions and interactions of various systems, and how to use this information to apply Newton's law to them. After such practice, students are much better prepared to use Newton's law to solve actual mechanics problems.

It is also important to teach explicitly the prerequisite knowledge required for description and interpretation. For example, one cannot describe the interactions of a system without an adequate knowledge of the properties of various forces. The properties of some of these (e.g., of contact forces like friction) are far from obvious and need to be spelled out more clearly than is commonly done.

Emphasizing both quantitative and qualitative descriptions. An excessive emphasis on mathematical formalism is prevalent in many physics courses—often leaving students with memorized formulas but little understanding. Occasionally, only some qualitative notions are introduced without ever being elaborated in more quantitative fashion. As the preceding paragraphs indicate, neither extreme reflects real scientific work. Thus it is desirable that instruction should strive for a more balanced complementary use of both quantitative and qualitative descriptions. This can be done in several ways.

(a) Embedding quantitative discussions in qualitative frameworks. It is useful to embed quantitative treatments in qualitative frameworks. For example, the properties of acceleration can be explored qualitatively for motions along straight and curved paths *before* deriving quantitative expressions for the acceleration in the special cases of linear or circular motion. As another example, *interaction* can be introduced as a general notion encompassing several quantitative concepts describing interactions (e.g., force, work, potential energy, torque). One can then point out that there exist important relations between motion and interactions, and these can ultimately be expressed in the form of quantitative mechanics principles (such as Newton's law or the law relating energy and work).

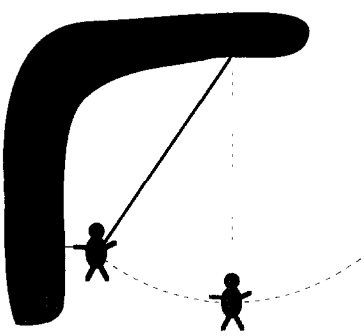

Fig. 8. A mountaineer, suspended from a rope attached to a ledge, holding on to a rock wall with a horizontal force.

(b) Solving qualitative as well as quantitative problems. Students can be given qualitative as well as quantitative problems—and the former may be equally challenging or instructive. For example, the problem illustrated in Fig. 8 asks students to determine whether the magnitude of the force exerted by the rope is larger than, smaller than, or equal to the mountaineer's weight (a) while he is at rest when holding on to the wall, (b) immediately after releasing his hold from the wall, and (c) when he is swinging past his lowest point where the rope is vertical. This mechanics problem is purely qualitative, yet it demands a good understanding of the relation between acceleration and forces.

(c) Qualitative checks and dependencies. Students should be encouraged to check their solutions of quantitative problems by assessing whether their results agree with qualitative predictions in special cases. If students are asked to express quantitative problem solutions in algebraic form, they can also readily explore how the results depend qualitatively on important parameters.

IV. KNOWLEDGE ORGANIZATION

The ability to use knowledge depends crucially on how well it is organized. For example, the folders in the file cabinets of one's office may be full of valuable information. But if the folders are haphazardly arranged, all this information is nearly useless because it is then almost impossible to find any specific information of interest. Although the information is potentially available, it is then not accessible. Further, the retrieval problem becomes increasingly severe if the information is more voluminous. (It is more difficult to find a needle in a larger haystack.)

Hence it is important that knowledge be effectively organized so as to facilitate the tasks of interest (e.g., selectively retrieving pertinent information, checking the consistency of the knowledge, generalizing it, augmenting it by further learning, etc.).

A. Common deficiencies

Characteristics of students' knowledge. Students' acquired scientific knowledge is often quite incoherent. Like much of everyday knowledge, it tends to be fragmented, consisting of separate knowledge elements that can often not be inferred from each other or from other knowledge.[16] For example, our observations[5] showed that students' knowledge of acceleration consists largely of miscellaneous bits of knowledge, often incorrect, which are unrelated to any more general conception.

This incoherence of students' knowledge was strikingly evident because students often encountered paradoxes between knowledge elements invoked by them—and were unable to resolve these paradoxes by any reasoning from more fundamental knowledge. For example, when trying to determine the acceleration of the pendulum bob at the extreme end of its swing [the point A in Fig. 2(c)], one student said:[5] "Acceleration is zero because it's not moving, so I'm sure of that. There's no acceleration on a non-moving object." But then the student went on to say: "Just because velocity at that point is zero, that doesn't mean there's no change in it, it's got to go from one direction to another." After going repeatedly back and forth between these two considerations, the student never could resolve this seeming paradox.

Many other such paradoxes reflect the incoherence of students' knowledge. For example, in considering the sled moving past the point A in Fig. 2(a), a student claimed that the sled's decreasing speed implies that the acceleration is opposite to the velocity. But then the student also thought that the acceleration ought to be vertically downward because of gravity. Again, the student was unable to resolve this apparent contradiction.

Experts' versus students' knowledge organizations. Many physicists proudly proclaim that physics, unlike organic chemistry or many other sciences, requires little memorization. However, students often complain that physics requires them to remember so many facts and formulas.

There is no real contradiction between these seemingly disparate perspectives. Good physicists have their knowledge organized in highly coherent form which makes it easy to remember and to infer much detailed information. But students' fragmented knowledge does not provide the benefits of such a coherent structure.

Inadequate instructional strategies. An incoherent knowledge organization, like that exhibited by these students, can be schematically represented by a set of disconnected knowledge elements like those indicated in Fig. 9(a). What are more effective forms of knowledge organization?

Some textbooks try to summarize physics knowledge by "formula lists."[7] But is such an undifferentiated laundry list of miscellaneous formulas a useful way to organize scientific knowledge?

Some teachers ask students to construct "concept maps" linking their concepts into connected networks of the form schematically indicated in Fig. 9(b). Such a network structure is certainly more coherent than a fragmented knowledge organization like that in Fig. 9(a). But does it significantly facilitate the selective retrieval of particular information? Indeed, suppose that one has a richly interconnected network of many knowledge elements. How then does one find one's way through such a jungle of interconnections to find any specific information of interest?

B. Cognitive considerations

Thus one needs to address the following central question: What kinds of knowledge organization can facilitate the retrieval of any particular information?

A hierarchical knowledge organization, like that schematically illustrated in Fig. 9(c), is well adapted to meet this requirement. In such a hierarchical structure, any knowledge element is elaborated into a few subordinate knowledge ele-

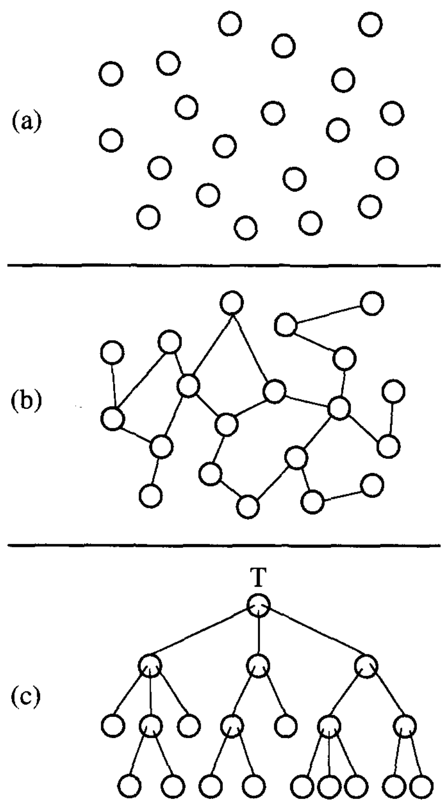

Fig. 9. Schematic representations of various knowledge organizations. (a) Incoherent knowledge consisting of largely disconnected knowledge elements. (b) Knowledge elements linked to form a network. (c) Hierarchical knowledge organization.

ments which can, in turn, be similarly elaborated. By means of such successive elaborations, any amount of detailed information can be incorporated without obscuring the main ideas. The hierarchical structure permits also a systematic and easily implemented retrieval process. Indeed, starting at the top level T of the structure, the retrieval process can be decomposed into successive stages each of which involves decisions among only a few alternatives. In this way one can efficiently search for a path leading to any information in the structure.

For example, geographical information about the United States is commonly organized hierarchically in the form of maps of different scales. In this organization the top level T is described by a coarse map of the entire United States. This map can then be elaborated into more detailed maps of the western, central, and eastern United States. Each of these maps can then be further elaborated into maps of the individual states in each of these regions. Each of these can then be further elaborated into still more detailed maps, and so forth. As we all know, such maps are very helpful in finding particular geographic information and in navigating throughout the entire United States.

There is also evidence indicating the efficacy of hierarchical organizations in scientific contexts. For example, a scientific argument or solution of a problem can be structured in different ways. It may be organized *linearly* as a sequence of 15 steps leading from the premises to the conclusion. Alternatively, it might be organized *hierarchically* into four major steps, each of which can be elaborated into three or four

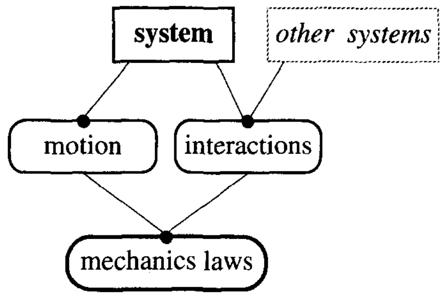

Fig. 10. Overview of mechanics.

detailed steps. In an experiment carried out by Eylon and myself some years ago,[18] we ensured that students learned the same argument organized either in linear or hierarchical form. Subsequent tests then showed that the students who had learned the argument in hierarchical form were much better able to remember the argument, to modify it, and to detect errors in it.

Knowledge about entire domains of physics can be hierarchically organized with some of the attendant benefits indicated by the preceding considerations. For example, mechanics can be described at its top level T by an overview like that in Fig. 10. This overview indicates that mechanics deals with the motions of systems and the interactions between them—and that it achieves its predictive power by mechanics laws specifying relations between motion and interactions.

It is then possible to elaborate hierarchically the knowledge in each of the preceding three categories, i.e., knowledge about the motion of a system (e.g., knowledge about velocities and accelerations), knowledge about interactions (e.g., knowledge about various long-range and contact forces), and knowledge about the relation between motion and interactions (e.g., knowledge about the laws of mechanics). For instance, this last knowledge can be elaborated into the three basic laws indicated in Fig. 11 (i.e., the laws of momentum, of angular momentum, and of energy). These laws can then be further elaborated by their special cases (e.g., by conservation laws valid under certain conditions).

Mechanics laws for a system

(Implied by Newton's law for a particle: $m\mathbf{a} = \mathbf{F}_{tot}$)

Momentum law $\quad \dfrac{d\mathbf{P}}{dt} = \mathbf{F}_{ext} \quad$ $m\mathbf{a} = \mathbf{F}_{tot}$ (for particle); $M\mathbf{A}_c = \mathbf{F}_{ext}$ (for CM)

Ang. mom. law $\quad \dfrac{d\mathbf{L}}{dt} = \tau_{ext} \quad$ $I\alpha = \tau_{ext}$ (if I constant)

Energy law $\quad \Delta E = W_{oth}$

Fig. 11. Basic laws of mechanics.

The preceding hierarchical organization is highly coherent. It does not involve a collection of miscellaneous facts and formulas, but only a few key ideas which can be easily remembered and flexibly elaborated. For example, an entire course in classical mechanics can be built around the three basic mechanics principles listed in Fig. 11.[19]

C. Instructional implications

Organizing a substantial body of knowledge effectively is not an easy task and beyond the capabilities of most inexperienced students. However, instruction can at least try to ensure (a) that students acquire knowledge which is in well-organized hierarchical form, and (b) that they can exploit such organization to help them remember and retrieve pertinent information.

The following means can help to achieve these goals.

Instead of presenting lots of facts and formulas, instruction should focus on only a few basic definitions and principles which can be systematically elaborated into more detailed knowledge when needed. Some of these basic ideas may usefully be described qualitatively, provided that they can be elaborated in more precise quantitative form when this is appropriate.

The organization of presented knowledge can be made more apparent by frequent summaries, as well as graphically by charts and diagrams.

Presenting knowledge in well-organized form is useful, but totally insufficient. The more important, and far more difficult, requirement is to ensure that the knowledge in students' heads is well organized. To this end, students must actively practice using well-organized knowledge. For example, they may be given a compact summary of basic knowledge—and then be asked to use this knowledge as the starting point for solving all problems and for justifying all their explanations.

V. PROBLEM SOLVING

A. Basic issues

The need for problem solving arises whenever one wants to attain desired goals. Problem solving is thus required for any purposeful behavior and is essential to attain the scientific goals of explaining, predicting, or designing.

Definition of a problem. Let me define more precisely what I mean by a problem so that the relevant issues can be clarified: A *problem* is a task that requires one to devise a sequence of actions leading from some initial situation to some specified goal. A well-specified sequence of such legitimate actions constitutes a *solution* of the problem.

A problem may be schematically represented by a diagram like that in Fig. 12(a). Here the point A represents the initial situation, the point G represents the goal, and a line linking two points represents a legitimate action leading from one situation to another. The highlighted sequence of links leading from A to G represents a solution of the problem.

Problems can be of widely different difficulties. Note that the difficulty of a problem depends not only on the problem itself, but also on the knowledge of the problem solver. (For example, a problem which is deemed difficult by one person may be quite easy for someone else who has solved similar problems before.)

Central difficulties of problem solving. Problem solving may often be intellectually demanding because it requires one to cope with the following two central difficulties.

(a)

(b)

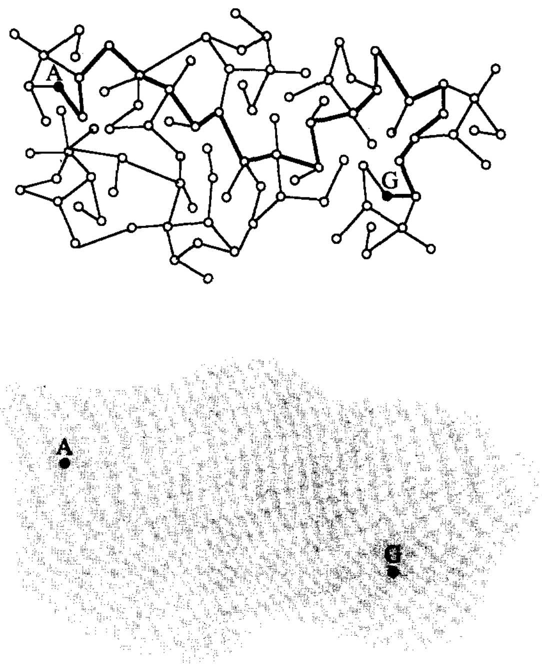

Fig. 12. (a) Schematic representation of a problem. (b) An unanalyzed problem.

(a) Decision making for search. As is apparent from Fig. 12(a), trying to solve a problem is somewhat like searching for a way through a maze. Starting from the initial situation A, there are many permissible actions that one can take [as indicated by the lines in Fig. 12(a)], but most cause one to get lost or stuck without ever getting to the goal G. Hence one faces the following difficulty: How can one make judicious decisions so as to choose, out of the very many possible action sequences which lead to nowhere, the one (or very few) which lead to the desired goal?

(b) Initial problem analysis. Before one can decide about alternative actions, one needs to address the following prior difficulty: How can one initially describe and analyze a problem so as to identify the useful possible actions among which one can choose? Indeed, a good initial analysis of a problem can greatly facilitate the task of finding its solution. Conversely, if the initial analysis of a problem is deficient, it may be impossible to find its solution despite all subsequent efforts. [Figure 12(b) illustrates schematically a poorly analyzed problem which does not even indicate the available pathways through the maze.]

B. Inadequacies of prevailing instruction

The most common method of teaching problem solving in physics relies largely on examples and practice. After some relevant physics knowledge has been presented, examples of some problem solutions are shown in the textbook or on the blackboard; then students are given practice solving other problems in homeworks.

The preceding method has the following limitations.

(a) Examples of problem solutions are *products* that reveal little about the process by which they were created. They may be judged for correctness, but they do not reveal how to make wise decisions so as to choose one principle rather than another, how to avoid solution paths which lead to nowhere, or how to recover from impasses when one is stuck. In other words, they do little to help students learn strategies for dealing with unfamiliar problems.

(b) Examples of problem solutions can lead to the misleading belief that the solution process resembles the product. But this is far from true! A computer program may consist of 2000 successive lines of code; but this does not mean that such a program is effectively constructed by writing successively line 1, then line 2, then line 3, and so forth until line 2000. Similarly, an English essay may consist of 500 successive sentences or the solution of a physics problem of 20 successive steps. But none of this implies that these products are constructed by the successive linear assembly of such ingredients. An effective process of generating these products may actually involve trial and error, planning, outlining, and progressive refinement.

(c) Adequate practice is certainly needed to learn problem solving. But it must be the *right kind* of practice! In athletics or music instruction, the wrong kind of practice cannot only be ineffective, but harmful (leading to poor habits which are difficult to break or even to injuries). Similarly, when students working on homework problems spend hours floundering, what they really practice is floundering. And if they spend most of their time haphazardly grabbing miscellaneous equations, they certainly do not practice valuable problem-solving skills.

It thus seems unwise to teach problem solving by relying excessively on examples and practice. Further, it is well known that students often have great difficulties learning how to solve physics problems, and that many students emerge from their physics courses with poor problem-solving skills very different from those of more expert problem solvers.[20,21] Thus the question arises: What might be more effective teaching methods?

C. Problem-solving method

The general approach outlined at the beginning of this paper suggests that one should try to analyze the thought processes required for effective problem solving, and that one should then try to teach explicitly a problem-solving method based on this analysis. No such method can be expected to guarantee the correct solution of every problem. It should, however, provide a heuristic strategy that substantially enhances the ability to solve problems and that is far more effective than the haphazard approaches adopted by most students.

Any method for problem solving needs to deal with the central problem-solving difficulties mentioned at the beginning of this section. Such a method must, therefore, specify how to analyze a problem initially and how to make judicious decisions in order to construct its solution. It must also specify how to check the solution so that it can be appropriately revised to ensure its correctness. The three major steps of such a problem-solving method are summarized in Fig. 13 and briefly discussed in the following paragraphs.

1. Initial problem analysis

The purpose of the initial problem analysis is to bring the problem into a form facilitating its subsequent solution.

To this end, one must first clearly specify the problem by describing the situation (with the aid of diagrams and useful symbols) and by summarizing the problem goals. This basic description of the problem can then be further refined by specifying the time sequence of events and by redescribing the situation in terms of more technical physics concepts (e.g., velocities, accelerations, and forces).

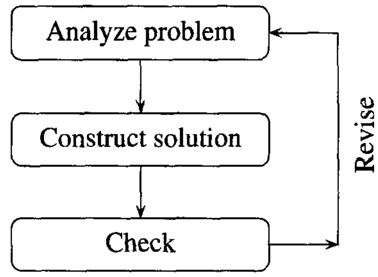

Fig. 13. Major steps of a systematic problem-solving method.

For example, the problem stated in Table I deals with a crate sliding down along a ramp. The initial analysis transforms this original problem statement into the problem description illustrated in Fig. 7(a). Note that this description summarizes the known information more transparently in visual form. Further, the goal of finding the "weight reading indicated by the scale" has been redescribed in terms of specific physics concepts, i.e., the goal is to find the *normal force* exerted on the ramp by the scale. (This redescription is not obvious to many students and makes the problem substantially easier.)

The initial analysis of a problem can greatly affect the ease of its solution. Conversely, deficiencies in an initial problem analysis are fatal; if they remain undetected, no amount of subsequent effort can lead to a correct solution of the problem. The initial analysis of a problem is thus quite important, but less simple than it may seem. Many students have difficulties reading a problem so as to extract the relevant information and visualize the situation. It is all too easy to ignore pertinent information or to make unwarranted assumptions.

2. Construction of a solution

Recursive problem decomposition. An effective strategy for constructing the solution of a problem is to "divide and conquer," i.e., to decompose the search for a solution into simpler subprocesses. The solution process involves thus repeated applications of the following two steps illustrated in Fig. 14: (a) Choosing some useful *subproblem* (i.e., some subsidiary problem facilitating the solution of the original problem), and (b) implementing the solution of this subproblem (or choosing a further subproblem if this is not possible). These steps can then be recursively repeated until the original problem has been solved.

Choice of subproblems. The decisions needed to solve a problem arise in choosing useful subproblems. Figure 14 indicates how such a choice can be made: One can (a) assess the status of the problem at any stage (by ascertaining what information is then known and the obstacles hindering a solution), (b) identify available options of subproblems likely to overcome the obstacles, and (c) select a useful option.

As indicated in Fig. 15, the two main obstacles (and corresponding remedial subproblems) that need to be addressed are the following.

(a) One may lack needed information. Correspondingly, one then needs to address the subproblem of finding a useful relation that can provide such information.

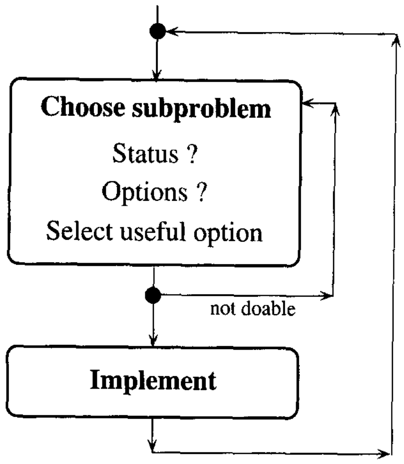

Fig. 14. Construction of a problem solution by recursive decomposition into subproblems.

(2) One may have available relations which are potentially useful, but contain undesirable features (e.g., unwanted unknown quantities). Correspondingly, one then needs to address the subproblem of eliminating the undesirable features. This can usually easily be done by using algebra to transform or combine available relations. (If not enough relations are available to do this, one is again faced with the obstacle of insufficient information.)

The most difficult task is the first of these, finding a useful relation to remedy the lack of needed information. As indicated in Fig. 15, this can be done by applying some *principle or definition* from one's knowledge base, to some particular *system* in the problem, at some specified *time* (or between some specified times). The decision difficulties have thus been reduced to easier separable choices of which principle, which system, and which time.

Knowledge organization facilitating choice. These choices are greatly facilitated if there are only very few reasonable options among which one needs to choose.

Indeed, suppose that one's knowledge base is effectively organized in hierarchical fashion. In a domain like mechanics, one then knows (as indicated in Fig. 10) that the central needed information is that relating motion and interactions—and that such information is provided by the three mechanics laws summarized in Fig. 11. Application of one of these three laws to a system in the problem should thus yield a useful relation for solving a mechanics problem. A choice

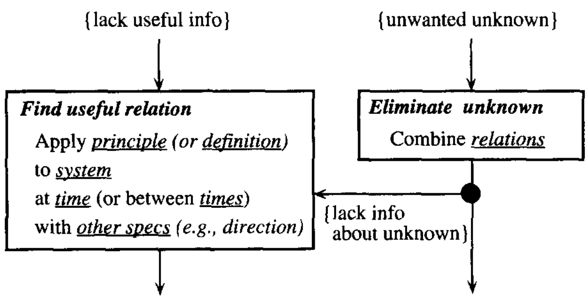

Fig. 15. Options for choosing useful subproblems.

among only three such alternative options is fairly easy (especially with ancillary knowledge about the conditions under which each of these laws is likely to be useful).

On the other hand, suppose that a student's physics knowledge consists of a large collection of miscellaneous formulas. Choosing a useful relation among all of these is then a difficult task that can often be unsuccessful.

The preceding comments indicate that an effective knowledge organization is of crucial importance in facilitating the decisions needed for problem solving.

Interpretation and description facilitating implementation. Once one has chosen to apply a particular principle to a particular system, this choice needs to be implemented. This implementation is, of course, greatly facilitated by all of one's previously acquired knowledge about interpretation and description (i.e., by the knowledge discussed in Secs. II and III).

Factors facilitating problem solving. To summarize, problem solving is facilitated by the following factors.

(1) The initial analysis of a problem identifies the important features of a problem and redescribes the problem in more useful form. [For example, the problem in Table I is transformed into the form of Fig. 7(a).]

(2) A search strategy, like that outlined in Figs. 14 and 15, helps to identify the kinds of choices that need to be made.

(3) A well-structured hierarchical knowledge organization facilitates these choices by reducing greatly the number of options that need to be considered. [For example, in mechanics the choice of useful principles is essentially reduced to one of the three laws listed in Fig. 11. In the case of the problem in Table I, the rather obvious choice among these is then the momentum law (or equivalently Newton's law) applied to the crate and to the ramp.]

(4) Implementation of any such choice is facilitated by previously acquired interpretation and description knowledge. [For example, in the case of the problem in Table I such description knowledge leads to the system diagrams of Fig. 7(b). Knowledge of how to interpret Newton's laws allows one then readily to write the equations resulting from application of this law to the crate and to the ramp.]

3. *Checking solutions*

It is essential to check any solution to assess whether it is correct and satisfactory—and to revise it appropriately if any deficiencies are detected. The following merely lists some standard checks applicable to any problem.

(a) Goals attained? (Has all wanted information been found?)

(b) Well specified? (Are answers expressed in terms of known quantities? Are units specified? Are both magnitudes and directions of vectors specified?)

(c) Self-consistent? (Are units in equations consistent? Are signs or directions on both sides of equations consistent?)

(d) Consistent with other information? (Are values sensible? Are answers consistent with special cases or with expected functional dependence? Are answers consistent with those obtained by another solution method?)

(e) Optimal? (Are answers and solution as clear and simple as possible? Are answers in general algebraic form?)

D. Instructional implications

Although the preceding comments about problem solving could be elaborated in much greater detail, they did touch on the central issues. Indeed, attempts to teach students effective problem-solving skills need deliberately to concentrate on only a few essential thought processes that students may actually learn to use. (It is wise to realize that less may be more. Much sage advice about problem solving can actually be of little practical value.)

One should certainly try to get students into the habit of decomposing the solution of every problem into the three previously discussed major stages (i.e., analyzing a problem, constructing its solution, and checking this solution). Correspondingly, any problem solution demonstrated by the instructor, or exhibited in the textbook, should similarly be decomposed into these three stages.

Initial problem analysis. When faced with a problem, most students try immediately to solve it without perceiving the need to analyze it carefully beforehand. Hence it is necessary to convince students that this behavior is often far less efficient than it naively seems, i.e., that time invested in the initial analysis of a problem is amply repaid by avoiding mistakes and the dangers of getting stuck. At least, students need to get in the habit of describing clearly the situation with appropriate diagrams and to specify clearly the problem goals.

It can help students to analyze some problems *without* actually solving them. This can convince students that such analysis is worthy of attention in its own right and that it is far from trivial (since students often make mistakes). Further, students can thereby practice separately skills of analysis and description which they will later need to deal with more complex problem-solving tasks.

Construction of solutions. It is useful to make decision processes highly explicit to students. Even the simple question "which principle could you apply to which system?" can help to dislodge a student from the paralysis of not knowing what to do next.

More generally, it helps to make explicit the available options among which students can choose. For example, one may give students a well-organized summary of *very few* basic relations (like the three mechanics laws in Fig. 11) and require that all problem solutions be based on these alone.

Students should be asked to include in their problem solutions sufficient explanatory comments to indicate their main decisions (e.g., which principle they applied to which system to find some new relation, or which available relations they combined to eliminate some unknown quantity). Doing this has two advantages: (a) Students are thereby encouraged to make explicit decisions based on solid knowledge, and are thus helped to avoid haphazard or intuitive guesses which are often wrong. (b) Instead of scribbling incomprehensible equations, students then produce recognizable solutions which can be understood and checked by others.

Checking solutions. Because they suffer from naive overconfidence or are reluctant to spend extra time, students often see little need to check their problem solutions. To remedy this situation, it helps to distinguish wrong answers from those which are *nonsensical* (because their incorrectness could readily be revealed by simple checks). One can then penalize more severely nonsensical answers than those which are merely wrong.

VI. PRACTICAL IMPLEMENTATION OF INSTRUCTION

The preceding sections identified some important thought processes needed to acquire flexibly usable physics knowledge. In particular, I discussed the need to interpret properly basic concepts and principles, to describe and organize scientific knowledge effectively, and to solve problems by methods facilitating their analysis and the decisions required for their solutions. I also pointed out some instructional implications and evidence that attention to such cognitive issues can lead to improved teaching effectiveness.

Much of my past work has been concerned with analyzing these issues separately and studying them in laboratory settings. But substantial improvements in teaching effectiveness can only be expected if all these cognitive issues are addressed *jointly* in actual instruction. Further, there is a large gap between small-scale studies and practical educational delivery. More recently I have, therefore, turned my attention to the following question: Can one design practical instruction which teaches explicitly some of the important thought processes identified in the preceding sections, and can one thereby help students to acquire more effectively usable scientific knowledge?

A. Special instructional materials

Accordingly, I have prepared instructional materials designed to achieve these aims in the context of an introductory college-level physics course for prospective science or engineering students. These materials, dealing with classical mechanics, consist of a text and a workbook closely coordinated to fulfill complementary functions.

The text is designed to present the basic ideas, and to be also useful for reference and review. It is fairly compact and coherently built around a small number of central ideas which are systematically elaborated. [These ideas are those summarized in Fig. 10, i.e., motion described by velocity and acceleration, a few important interactions, and the relation between motion and interactions (ultimately elaborated into the three mechanics laws of Fig. 11).] The text deliberately emphasizes reasoning methods, i.e., methods for interpreting important concepts and principles, methods for describing situations both qualitatively and quantitatively, and methods of problem solving. It also tries to sequence the instruction so that concepts or principles (like acceleration or Newton's law) are discussed qualitatively before they are elaborated in more quantitative detail.

The workbook is designed to engage students actively in their learning (since mere study of the text would scarcely lead to any learning at all). Thus almost every section of the text is accompanied by a corresponding workbook section which engages students in the active exploration and interpretation of any new ideas introduced in the text. The workbook also provides students with practice in applying systematic methods to solve various qualitative and quantitative problems. Finally, the workbook contains hints and answers so as to provide a fair amount of guidance and feedback. It thus tries to help students learn without excessive needs to rely on outside assistance.

Planned publication by Wiley should make these materials available for wider distribution.[22]

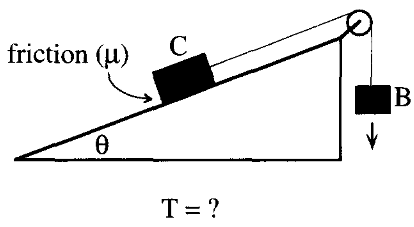

Fig. 16. One of the problems on the final examination.

B. Fostering active student learning

Even the best instructional materials are worthless unless they are properly used. In particular, no learning can occur unless students engage in active thinking. (Physics knowledge cannot yet be injected like a vaccine.) Indeed, active engagement of students in their learning is a feature common to many recent instructional innovations (like those described in the Refs. 23–31).

Accordingly, I have tried to use my instructional materials in the context of a course structure designed to promote active student involvement. One important means has been the extensive use of the workbook and of homework assignments taken from it. Other means have involved making lectures more interactive, structuring discussion sections so that students work collaboratively in small groups, and giving frequent diagnostic tests.

For historical and logistical reasons, the introductory physics course at Carnegie–Mellon University has no laboratory associated with it. Otherwise, some of my past work[32] suggests that the laboratory too could help to teach students effective scientific thinking skills.

C. Evidence of instructional efficacy

Since the preceding work is still in progress, there is still no definitive evidence about the practical effectiveness of these instructional materials and methods. However, I have been able to gather some preliminary data.

The first semester of the introductory physics course at Carnegie–Mellon University has recently been taught by two different professors in two simultaneous sections covering the same topics. (The students have been randomly assigned to these sections.) In the fall of 1991, and again in the fall of 1992, one of these sections was taught by me, using my own materials and approach. The other section was taught by another experienced professor along more conventional lines.[33]

My own instructional materials at those times were still preliminary and incomplete, covering only about half the topics discussed in the course. Nevertheless, the existence of these simultaneous sections permitted some detailed comparisons, particularly on common questions on the final examinations.

Comparative performance on final examinations. For example, Fig. 16 illustrates one of the common questions on the final examination given in 1992. Students were asked to find the magnitude of the tension force exerted by the string when the block B descends so that the crate C moves up along the ramp. In my "special" section, 70% of the students reasoned correctly to solve this problem and obtained the correct answer (except possibly for minor algebraic mistakes). In the other "ordinary" section only 10% of the students were equally successful.

We analyzed four such common problems on the final examination. If one adopts the stringent criterion that the solution to any such problem is considered correct only if *all* its parts have been correctly answered, then correct solutions were (on the average) obtained by about 45% of students in the special section and by about 10% percent of students in the ordinary section. The performance of students in the special section, although certainly far from perfect, was thus considerably better than that of students in the ordinary section.

Students' performance on the common questions of the earlier 1991 final examination was similarly consistently in favor of the special section.

Reduced prevalence of misconceptions. It was particularly interesting to analyze the kinds of mistakes and misconceptions exhibited by the students. The relevant data could be obtained not only from answers to examination questions, but also from detailed videotaped observations of individual students asked to deal with various physics problems. These data indicate that students enrolled in the special section were much less prone to display commonly occurring misconceptions or scientific nonsense.

For example, in the problem illustrated in Fig. 16, 35% of the students in the ordinary section claimed that the magnitude of the tension force exerted by the string is simply equal to the weight of block B (thus displaying a common misconception and totally ignoring the acceleration of this block). By contrast, only 7% of students in the special section made this mistake. Similarly, 35% of the students in the ordinary section did not properly apply Newton's law, but simply claimed that the sum of all forces on the crate is zero. Only 7% of the students in the special section made this mistake.

Usability by other instructors. In the fall of 1993, the professor who had taught the ordinary section during the preceding year asked to use my materials when teaching the course again. This collaborative arrangement caused me to lose opportunities for making comparisons, but did show that my instructional materials and methods could also be used effectively by someone other than myself.[34]

Encouraging indications. The preceding results encourage the belief that the instructional approach discussed in the preceding pages has some of the desired beneficial effects and may be translated into practical classroom instruction.

VII. IMPLEMENTATION DIFFICULTIES

Practical implementations are particularly difficult when they depend on the cooperation of other people. For then they require not only the solutions of all intrinsic problems, but also appropriate beliefs and actions on the part of the pertinent people.

For example, consider a hypothetical situation where medical science has succeeded in identifying the causes of most diseases and in producing pills that can reliably cure all of them. Suppose, however, that people do not heed scientific medicine because they believe in folk medicine or natural healing. Or suppose that they do not take the pills according to the recommended regimens. Then the health of the population will remain poor despite all medical knowledge, i.e., practical medical implementation will have failed.

As pointed out in the following paragraphs, practical educational implementations face similar difficulties (quite apart from motivational factors). For example, suppose that we understood perfectly the thought and learning processes needed for physics. All this understanding would still be insufficient to ensure practical educational implementation if students have misleading beliefs about science or do not actually engage in the recommended learning activities.

A. Changing students' naive notions about science

It is well known that students come to the study of science with many naive prior notions about the physical world and that these notions are very difficult to change.[1–4] It is less well appreciated that students also come with many naive conceptions about the goals of science and about the kinds of thinking needed for science. These conceptions, imported from everyday life or from prior schooling, are even more difficult to change than students' naive notions about the physical world. Further, their effects are all pervasive, affecting greatly what students try to learn and how they go about learning it.[35,36]

Examples of student beliefs. For example, many students view science as a valuable collection of facts and formulas–and thus pursue the goal of memorizing these. On the other hand, most physicists view science as a small body of basic knowledge enabling wide-ranging inferences about many observable phenomena. When physicists try to teach reasoning abilities needed to make scientific inferences, they pursue then a learning goal very different from that envisaged by the students.

As another example, many students come from high-school courses that emphasized formulas and number crunching—and thus believe that these are the important activities needed for doing science. Hence they resist learning to do qualitative reasoning or to obtain algebraic results that can reveal important qualitative relationships.

Many other examples of such student beliefs could be listed.[35–38] But even the few examples just mentioned are sufficient to indicate that such beliefs can greatly influence students' reactions to instruction and the extent to which they can benefit from it. What is even worse, valuable instructional innovations can easily be resented when they are not in accord with students' naive perspectives. (For instance, such student resentments have cropped up in recent innovations in calculus instruction.[39])

Instructional implications. An introductory physics course needs thus also to discuss explicitly the goals of science and the ways of thinking useful in science. These issues cannot merely be addressed by a few occasional remarks. They need to be constantly kept in students' focus, and be used as a framework within which more specific scientific knowledge and methods are embedded.

B. Providing adequate guidance and feedback

Even the best instructional materials and methods are useless if students do not actually engage in the recommended learning activities. This is well recognized in efforts designed to train good atheletes or musical performers. There coaches or teachers provide very frequent supervision, with the guidance and feedback necessary to ensure that students acquire good habits—and to prevent bad habits which may be difficult to break or even lead to injuries.[40]

Inadequacies of guidance and feedback. The needs for adequate supervision are even greater when one is trying to teach students the intellectual abilities needed to deal with science. Unfortunately, science classes usually provide far too little guidance and feedback to ensure that students actually engage is useful learning activities.

For example, the instructor and textbook may discuss effective problem-solving methods, demonstrate these in class, give students homework to practice these, and correct these homeworks afterwards. But many students, left to their own devices, still do homework problems by reaching haphazardly for various formulas—and hours of such practice merely perpetuate bad habits. All the preceding instructional efforts may, therefore, be far less effective than an hour of individual tutoring by the instructor—because such tutoring provides much better guidance and feedback.

Instructional implications. How then can one provide students with adequate individual guidance and feedback in practical contexts dealing with many students? In my judgment, this is a fundamentally important problem which, if left unsolved, will remain a bottleneck hindering the implementation of even the best-designed instruction. Computers acting as tutors might help appreciably. Collaborative learning could also help (provided that students are specifically trained to engage in this effectively). But all such suggestions are difficult to implement well in cost-effective ways. Thus there remains a real challenge.

VIII. CONCLUDING REMARKS

All too often introductory physics courses "cover" numerous topics, but the knowledge actually acquired by students is nominal rather than functional. If students are to acquire basic physics knowledge that they can flexibly use, it is necessary to understand better the requisite thought processes and to teach these more explicitly. In the preceding pages I have attempted to identify and briefly analyze some of these thought processes. In particular, these include those needed to interpret basic concepts and principles, to describe and organize knowledge effectively, and to solve problems (i.e., to analyze problems, to construct their solutions by judicious decisions, and to check these solutions).

I have also indicated how these thought processes may be taught. Indeed, I believe that it is necessary to teach these deliberately if one wants to improve significantly students' learning of physics. However, even such explicit teaching may, in practice, be insufficient to achieve the desired instructional goals. It is also necessary to modify students' naive notions about the nature of science, and to provide better guidance and feedback to individual students in large classes. Thus much remains to be done by scientists engaged in systematic efforts to improve instruction. It is my hope that some of you will help contribute to this endeavor.

ACKNOWLEDGMENTS

The National Science Foundation has partially supported a substantial fraction of the work summarized in this paper, most recently through Grant No. MDR-9150008. Many co-workers have contributed to this work over the years. These individuals include Jill Larkin and Bat-Sheva Eylon; also Joan Heller, Lisa Quinn, Peter Labudde, Sue Allen, Lisa Scott, and Molly Johnson. Bruce Sherwood and Ruth Chabay have been good colleagues at Carnegie–Mellon University. Herbert Simon, long before I knew him personally, has helped by legitimizing my own interests in cognitive processes and, more recently, by his active support of educational applications. Professor Brian Quinn allowed me to

compare his section of the Carnegie–Mellon physics course with my own. Professor Gregg Franklin, after doing likewise, later also collaborated with me and used my instructional materials in his own course.

[1]L. C. McDermott, "Research on conceptual understanding in mechanics," Phys. Today **37**, 24–32 (July 1984).

[2]I. A. Halloun and D. Hestenes, "The initial knowledge state of college students," Am J. Phys. **53**, 1043–1055 (1985).

[3]I. A. Halloun and D. Hestenes, "Common sense concepts about motion," Am. J. Phys. **53**, 1056–1065 (1985).

[4]L. C. McDermott and P. S. Shaffer, "Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding," Am. J. Phys. **60**, 994–1003 (1992).

[5]F. Reif and S. Allen, "Cognition for interpreting scientific concepts: A study of acceleration," Cogn. Instruct. **9**, 1–44 (1992).

[6]L. C. McDermott and P. S. Shaffer, "Research as a guide for improving instruction in introductory physics," Ph. D. dissertation, Department of Physics, University of Washington, 1993 (unpublished), Chap. 4.

[7]F. Reif, "Interpretation of scientific or mathematical concepts: Cognitive issues and instructional implications," Cogn. Sci. **11**, 395–416 (1987).

[8]P. Labudde, F. Reif, and L. Quinn, "Facilitation of scientific concept learning by interpretation procedures and diagnosis," Int. J. Sci. Ed. **10**, 81–98 (1988).

[9]F. Reif, "Instructional design, cognition, and technology: Applications to the teaching of scientific concepts," J. Res. Sci. Teach. **24** 309–324 (1987).

[10]J. I. Heller and F. Reif, "Prescribing effective human problem-solving processes: Problem description in physics," Cog. Instruct. **1**, 177–216 (1984).

[11]L. C. McDermott (private communication).

[12]A. Einstein, *Ideas and Opinions* (Crown, New York, 1954), p. 225.

[13]J. Hadamard, *The Psychology of Invention in the Mathematical Field* (Princeton University Press, Princeton, NJ, 1945) (reprinted by Dover Publications, New York, 1954, pp. 147–148).

[14]J. Bernstein, "Master of the trade," The New Yorker, p. 84 (3 December 1979).

[15]James Gleick, *Genius: The Life and Science of Richard Feynman* (Pantheon Books, New York, 1992), p. 244.

[16]A diSessa, "Knowledge in pieces," *Constructivism in the Computer Age*, edited by G. Forman and P. B. Pufall (Lawrence Erlbaum, Hillsdale, NJ, 1988), pp. 49–70.

[17]See, for example, H. C. Ohanian, *Physics* (Norton, New York, 1985), p. 563.

[18]B. Eylon and F. Reif, "Effects of knowledge organization on task performance," Cogn. Instruct. **1**, 5–44 (1984).

[19]This hierarchical structure of mechanics remains pertinent in more advanced courses where the mechanics laws would also include Lagrangian and Hamiltonian formulations.

[20]J. H. Larkin, J. McDermott, D. P. Simon, and H. A. Simon, "Expert and novice performance in solving physics problems," Science **208**, 1335–1342 (1980).

[21]M. T. H. Chi, P. J. Feltovich, and R. Glaser, "Categorization and representation of physics problems by experts and novices," Cogn. Sci. **5**, 121–152 (1981).

[22]F. Reif, *Understanding Basic Mechanics* (Wiley, New York, 1995).

[23]A. B. Arons, *A Guide to Introductory Physics Teaching* (Wiley, New York, 1990).

[24]R. R. Hake, "Socratic pedagogy in the introductory physics laboratory," Phys. Teach. **30**, 546–552 (1992).

[25]P. Heller, R. Keith, and S. Anderson, "Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving," Am. J. Phys. **60**, 627–636 (1992).

[26]P. Heller and M. Hollabaugh, "Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups," Am. J. Phys. **60**, 637–644 (1992).

[27]P. W. Laws, "Calculus-based physics without lectures," Phys. Today **44**, 24–31 (December 1991).

[28]E. Mazur, *Peer Instruction: A User's Manual* (Harvard University Press, Cambridge, MA, 1993).

[29]L. C. McDermott, "Millikan lecture 1990: What we teach and what is learned-closing the gap," Am. J. Phys. **59**, 301–315 (1991).

[30]A. Van Heuvelen, "Learning to think like a physicist: A review of research-based instructional strategies," Am. J. Phys. **59**, 891–897, (1991).

[31]A. Van Heuvelen, "Overview, case study physics," Am. J. Phys. **59**, 898–907 (1991).

[32]F. Reif and M. St. John, "Teaching physicists' thinking skills in the laboratory," Am J. Phys. **47**, 950–957 (1979).

[33]This other section used the textbook by H. D. Young, *Physics, 8th ed.* (Addison-Wesley, Reading, MA, 1992); the 7th edition of this book was used in the first of these courses.

[34]The materials have recently also been used by Professor Richard Hake at Indiana University, as well by some other instructors in Pittsburgh.

[35]F. Reif and J. H. Larkin, "Cognition in scientific and everyday domains: Comparison and learning implications," J. Res. Sci. Teach. **28**, 733–760 (1991).

[36]A. H. Schoenfeld, "Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance," Cogn. Sci. **7**, 329–363 (1983).

[37]D. Hammer, "Epistemological beliefs in introductory physics," Cogn. Instruct. **12**, 151–183 (1994).

[38]W. M. Roth and A. Roychoudhury, "Physics students' epistemologies and view about knowing and learning," J. Res. Sci. Teach. **31**, 5–30 (1994).

[39]E. Culotta, "The calculus of education reform," Science **255**, 1060–1062 (1992).

[40]For example, Jascha Heifetz, the famous violinist, was initially taught by his father. The latter deemed carefully supervised practice so important that he was always present when Jascha practiced, and locked up the violin to prevent Jascha from practicing in his absence. [A. Benoist, *The Accompanist: An autobiography of Andre Benoist* (Paganiniana, Neptune, NJ, 1978), pp. 287 and 284].

UNITS: THE MOONEY UNIT

The plasticity of raw, or unvulcanized, rubber is sometimes given in terms of the torque on a disc situated in a cylindrical vessel containing rubber at a temperature of 100 °C. The vessel rotates at 2 revolutions per minute and the torque on the disc after the vessel has been rotating for T minutes is measured on an arbitrary scale calibrated between 0 and 200. The number of the scale indicates the plasticity of the rubber in Mooney units, the result being expressed as so many Mooney units in T minutes. The unit is named after Dr Melvin Mooney, who devised the method in 1934.

H. G. Jerrard and D. B. McNeill, *A Dictionary of Scientific Units Including Dimensionless Numbers and Scales* (Chapman and Hall, London and New York, 1980), p. 90.