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

$$S_i \to S_f. \quad (1)$$

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 $\to$). (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 $S_i \to S_f$ will lead me to do the following: (a) Indicate some common inadequacies of students' initial abilities. (b) Analyze the thought processes re-

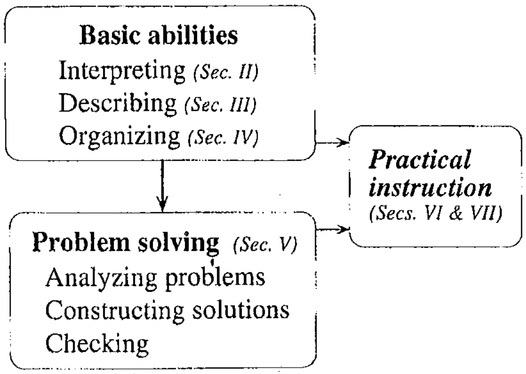

Fig. 1. Central cognitive issues important for scientific work.

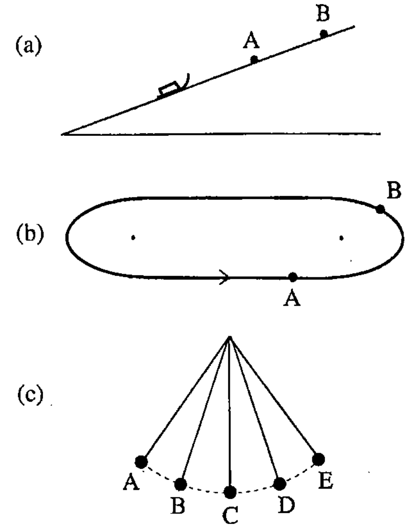

Fig. 2. Situations used for testing the interpretation of acceleration. (a) Sled sliding along a hill. (b) Car traveling around a horizontal racetrack. (c) Oscillating pendulum bob.

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.

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

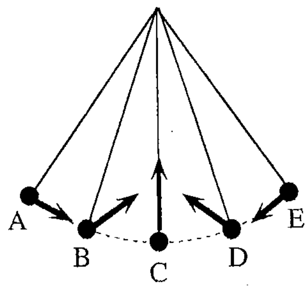

Fig. 3. Accelerations of a pendulum bob.

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."

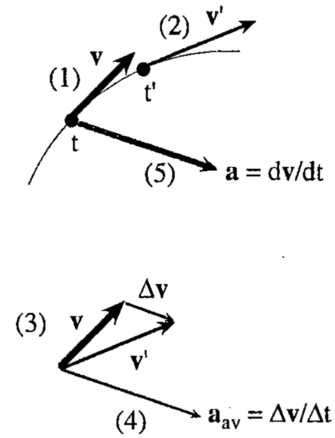

Fig. 4. Defining method for acceleration.

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°.

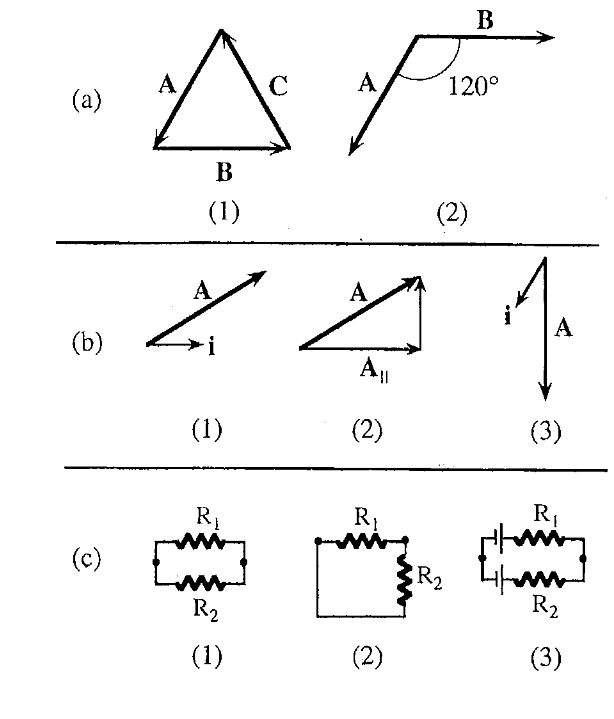

Fig. 5. Fallibility of interpretation by recognition. (a) Angle between two vectors. (b) Component vector of a vector. (c) Resistors connected in parallel.

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 $\mathbf{a}$ 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.

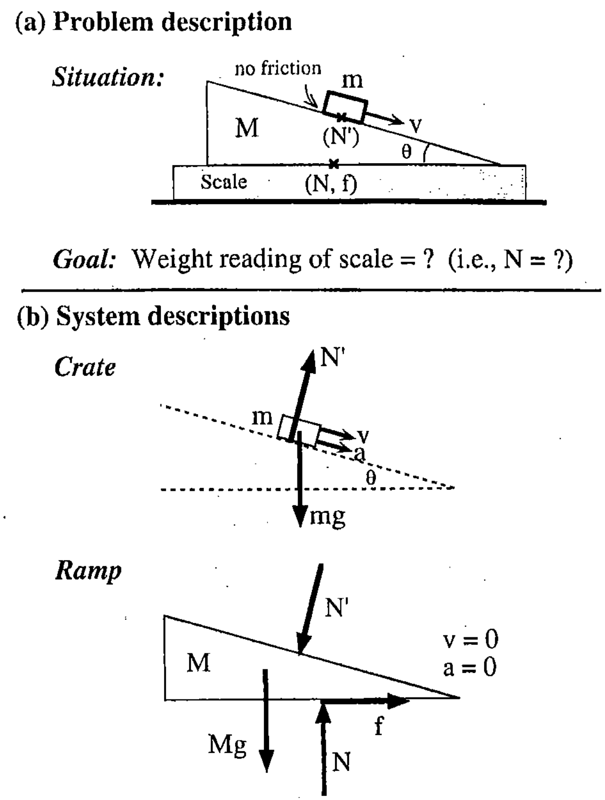

Fig. 7. (a) Useful description of the problem in Table I. (b) System diagrams describing the crate and the ramp (with vectors not yet decomposed into convenient components).

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).

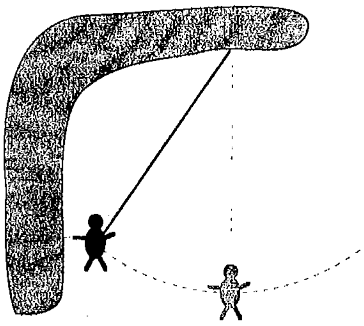

Fig. 8. A mountaineer, suspended from a rope attached to a ledge, holding on to a rock wall with a horizontal force.

(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-

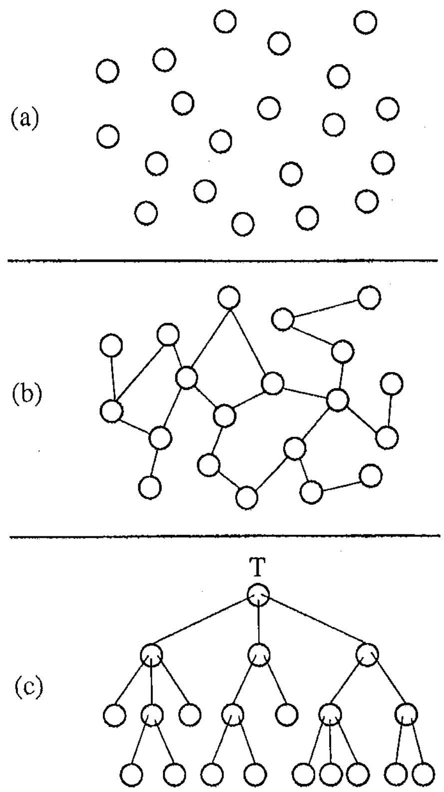

Fig. 9. Schematic representations of various knowledge organizations. (a) Incoherent knowledge consisting of largely disconnected knowledge elements. (b) Knowledge elements linked to form a network. (c) Hierarchical knowledge organization.

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 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.

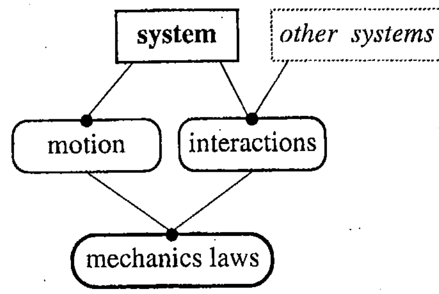

Fig. 10. Overview of mechanics.

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	$\frac{d\mathbf{P}}{dt} = \mathbf{F}_{ext}$	$m\mathbf{a} = \mathbf{F}_{tot}$ (for particle) $M\mathbf{A}_c = \mathbf{F}_{ext}$ (for CM)
Ang. mom. law	$\frac{d\mathbf{L}}{dt} = \tau_{ext}$	$I\alpha = \tau_{ext}$ (if I constant)
Energy law	$\Delta E = W_{oth}$	

Fig. 11. Basic laws of mechanics.