Categorization and Representation of Physics Problems by Experts and Novices*

MICHELENE T. H. CHI

PAUL J. FELTOVICH

ROBERT GLASER

University of Pittsburgh

The representation of physics problems in relation to the organization of physics knowledge is investigated in experts and novices. Four experiments examine (a) the existence of problem categories as a basis for representation; (b) differences in the categories used by experts and novices; (c) differences in the knowledge associated with the categories; and (d) features in the problems that contribute to problem categorization and representation. Results from sorting tasks and protocols reveal that experts and novices begin their problem representations with specifiably different problem categories, and completion of the representations depends on the knowledge associated with the categories. For, the experts initially abstract physics principles to approach and solve a problem representation, whereas novices base their representation and approaches on the problem's literal features.

CATEGORIZATION AND REPRESENTATION OF PHYSICS PROBLEMS BY EXPERTS AND NOVICES

This paper presents studies designed to examine differences in the ways expert and novice problem solvers represent physics problems and to investigate implications of these differences for problem solution. A *problem representation* is a

*This research program, conducted at the Learning Research and Development Center, is supported in part by contract No. N00014-78-C-0375, NR 157-421 of the Office of Naval Research, and in part by the National Institute of Education. Portions of this paper were presented by the first author at the meeting of the American Educational Research Association, San Francisco, April 1979, and at the meeting of the Psychonomic Society, Phoenix, November 1979. The authors are grateful for the help of Andrew Judkis, Ted Rees, and Christopher Roth, for comments, data collection, analysis, and editing. We particularly appreciate the generosity of the physics professors and graduate students who contributed their time, especially Ned S. VanderVen. Jill Larkin deserves special thanks for her contributions and insightful comments to Study Four. Reprint requests should be sent to Michelene Chi, Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pa. 15260.

cognitive structure corresponding to a problem, constructed by a solver on the basis of his domain-related knowledge and its organization. A representation can take a variety of forms. Greeno (1977), for example, has proposed the representation of a problem as a constructed semantic network containing various components. Some of these correspond closely with the problem as stated, including the initial state (i.e., the "givens"), the desired goal, and the legal problem-solving operators (Newell & Simon, 1972). In addition, a representation can contain embellishments, inferences, and abstractions (Heller & Greeno, 1979). Since embellishment is one way of judging a solver's "understanding" of a problem (Greeno, 1977), it is possible that with increasing experience in a domain, the representation becomes more enriched. The research described here explores the changes in problem representation that emerge as a result of developing subject-matter expertise.

It is well known by now that the quality of a problem representation influences the ease with which a problem can be solved (Hayes & Simon, 1976; Newell & Simon, 1972). In physics, Simon and Simon (1978) have attributed the expert's "physical intuition" to the quality of the problem representation. The current consensus is that the expert's representation is superior because it contains a great deal of "qualitative" knowledge. De Kleer (1977), for example, has introduced both "quantitative" and "qualitative" components in the expert's representation of a physics problem where the qualitative component includes nonmathematical semantic descriptions of physical objects and their interactions. Novak's (1977) program ISSAC also suggested some characteristics of qualitative representation. In this program, physical objects from a problem statement are represented not literally, but rather, as abstract object categories—canonical object frames—each of which serves an equivalent physics role (e.g., pivot, lever, or point mass). The *canonical object frame* is a knowledge structure that augments the information about an object stated in a problem with associated information from the knowledge base. In his later work, Novak has proposed the inclusion of problem types in categorization by types (Novak & Araya, 1980). Categorization of a problem as a type would cue associated information in the knowledge base. Similarly, Reif (1979) has proposed a problem-solving model in which an initial step is a representation or "redescription of any problem in terms of concepts provided by the knowledge base" (p. 1). This knowledge base is arranged around "problem schemata," each of which contains information necessary to solve a specific category of problems.

The hypothesis guiding the present research is that the representation is constructed in the context of the knowledge available for a particular type of problem. The knowledge useful for a particular problem is indexed when a given physics problem is categorized as a specific type. Thus, expert-novice differences may be related to poorly formed, qualitatively different, or nonexistent categories in the novice representation. In general, this hypothesis is consistent with the "perceptual chunking" hypothesis for experts (e.g., Chase & Simon,

1973) and its more general cognitive ramifications (e.g., Chase & Chi, in press), which suggest that much of expert power lies in the expert's ability to quickly establish correspondence between externally presented events and internal models for these events.

More particularly, some evidence already exists in the literature to suggest that solvers represent problems by category and that these categories may direct problem solving. First, Hinsley, Hayes, and Simon (1978) found that college students can categorize algebra word problems into types, and that this categorization can occur very quickly, sometimes even after reading just the first phrase of the problem statement. For example, if subjects were to hear the words "a river steamer," then they might surmise that the problem was one about current, perhaps comparing the rates of going upstream and downstream. The ability to categorize problems quickly suggested to Hinsley et al. (1978) that "problem schemata" exist and can be viewed as interrelated sets of knowledge that unify superficially disparate problems by some underlying features. Secondly, in chess research, it appears that experts' superiority in memorizing chessboard positions arises from the existence of a large store of intact and well-organized chess configurations or patterns in memory (Chase & Simon, 1973). It is plausible that a choice among chess moves (analogous to physics solution methods) results from a direct association between move sequences and a configural (chunked) representation of the surface features of the board. Finally, from research in medical diagnosis, there is evidence to suggest that expert diagnosticians represent particular cases by general categories, and that these categories facilitate the formation of hypotheses during diagnosis (Pople, 1977; Wortman, 1972).

The accumulation of evidence for the importance of categorization in expert problem solving leads us to examine the role of categorization in expert physics problem solving: particularly, to investigate the relationships between such categorization and subsequent attempts at solution. The following series of studies attempts to determine: the categories that experts and novices impose on physics problems (Studies One and Two); the knowledge which these categorical representations activate in the problem solver (Study Three); and the cues or features of problems which subjects use to choose among alternative categories (Study Four).

Study One: Problem Sorting

The objective of the first study was to determine the kinds of categories subjects (of different experience) impose on problems. Using a sorting procedure, we asked eight advanced PhD students from the physics department (experts) and eight undergraduates (novices) who had just completed a semester of mechanics, to categorize 24 problems selected from Halliday and Resnick's (1974) *Fundamentals of Physics*, beginning with Chapter 5, Particle Dynamics, and ending with Chapter 12, Equilibrium of Bodies. Three problems were selected from each chapter, and these were individually typed on 3 × 5 cards. Instructions were to sort the 24 problems into groups based on similarities of solution. The subjects were not allowed to use pencil and paper and, thus, could not actually solve the problems in order to sort them. As a test of consistency, subjects were asked to re-sort the problems after the first trial. Following this, they were asked to explain the reasons for their groupings. The time taken to sort on each trial was also measured.

Analysis of Gross Quantitative Results

No gross quantitative differences between the sorts produced by the two skill groups were observed. There were no differences in the number of categories produced by each group (8.4 for the experts and 8.6 for the novices), and the four largest categories produced by each subject captured the majority of the problems (80 percent for the experts and 74 percent for the novices). Likewise, experts and novices were equally able to achieve a stable sort within the two trials, that is, their second sort matched their first sort very closely. This suggests that their sorting pattern was not ad hoc, but rather, was based on some meaningful representation.

There were, however, some differences in the amount of time it took experts and novices to sort the problems. In fact, experts took longer (18 minutes or 45 seconds per problem, on the average) to sort the problems in the first trial than novices (12 minutes or 30 seconds). Both groups were relatively fast at sorting the second trial (4.6 minutes for the experts and 5.5 minutes for the novices). The speed with which the problems were sorted on the second trial (about 12 seconds per problem) suggests that subjects probably did not have to go through the entire process of "understanding" each problem again. Since the problems were all categorized after the first trial, the subjects probably needed only to identify the cues that elicited category membership.

In general, these quantitative data suggest that both experts and novices were able to categorize problems into groups in a meaningful way. Other than the difference in the time taken to sort on the first trial, there was little difference between skill groups. The critical question then becomes: what are the bases on which experts and novices categorize these problems?

Qualitative Analysis of the Categories

Analyses of Four Pairs of Problems. A cluster analysis (Diameter method) was performed on the problems grouped together by the experts and those by the novices. Such an analysis shows the degree to which subjects of each skill group agree that certain problems belong to the same group. One way to interpret the cluster analysis is to examine only those problems that were grouped together with the highest degree of agreement among subjects.

Our initial analysis centered on four pairs of problems. Figures 1 and 2 contain the diagrams of pairs of problems that were grouped together by the novices and the experts, respectively. These diagrams can be drawn to depict the physical situations described in the problem statements, and are sometimes given along with a problem statement (although no diagrams were given to the subjects in our studies). All eight novices grouped the top pair (Figure 1) together, and seven of the eight novices grouped the bottom pair. Both pairs of problems (Figure 2) were grouped together by six of the eight experts.

Examination of the novice pairs (Figure 1) reveals certain similarities in the surface structures of the problems. By "surface structures," we mean: (a) the objects referred to in the problem (e.g., a spring, an inclined plane); (b) the literal physics terms mentioned in the problem (e.g., friction, center of mass); or (c) the physical configuration described in the problem (i.e., relations among physical objects such as a block on an inclined plane). Each pair of problems in Figure 1 contains the same object components and configurations—circular disks in the upper pair, blocks on an inclined plane in the lower pair.

The suggestion that novices categorize by surface structure can be confirmed by examining subjects' verbal descriptions of their categories. (Samples are given in the figures.) Basically, according to their explanations, the top pair of problems involves "rotational things" and the bottom two problems "blocks on inclined planes."

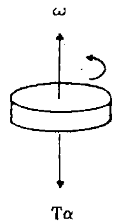
To reiterate, the novices' use of surface features may involve either keywords given in the problem statement or abstracted visual configurations, that is, the presence of identical keywords (such as friction) is one criterion by which novices group problems as similar. Yet, novices were also capable of going beyond the word level to classify by types of physical objects. For example, "merry-go-round" and "rotating disk" are classified as the same object, as is the case for the top pair of problems in Figure 1.

For experts, surface features do not seem to be the bases for categorization. There is neither great similarity in the keywords used in the problem statements, nor visual similarity apparent in the diagrams depictable from each pair of problems shown in Figure 2. Nor is the superficial appearance of the equations that can be used on these problems the same. Only a physicist can detect the similarity underlying the expert's categorization. It appears that the experts classify according to the major physics principle governing the solution of each problem. The top pair of problems in Figure 2 can be solved by applying the Law of Conservation of Energy; the bottom pair is better solved by applying Newton's Second Law ($F = MA$). The verbal justification of the expert subjects confirms this analysis. If "deep structure" is defined as the underlying physics law applicable to a problem; then, clearly, this deep structure is the basis by which experts group the problems.

Analysis of Categories. Further insight into the ways subjects categorize problems is given by the descriptions subjects gave for the categories they

Diagrams Depicted from Problems Categorized by Novices within the Same Groups	*Novices' Explanations for Their Similarity Groupings*

Problem 10 (11)

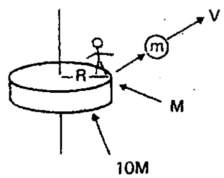


Novice 2: "*Angular* velocity, *momentum*, circular things"

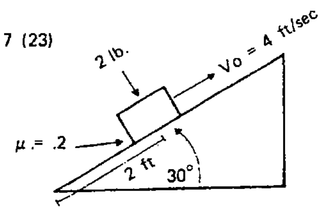
Novice 3: "*Rotational* kinematics, *angular* speeds, *angular* velocities"

Novice 6: "Problems that have something *rotating; angular* speed"

Problem 11 (39)



Problem 7 (23)



Novice 1: "These deal with blocks on an *incline plane*"

Novice 5: "*Inclined plane* problems, coefficient of *friction*"

Novice 6: "Blocks on *inclined planes* with angles"

Problem 7 (35)

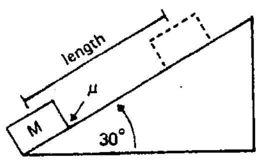
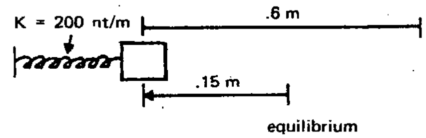


Figure 1. Diagrams depicted from two pairs of problems categorized by novices as similar and samples of three novices' explanations for their similarity are provided. Problem numbers given represent chapter, followed by problem number from Halliday and Resnick (1974).

Diagrams Depicted from Problems Catergorized by Experts within the Same Groups	*Experts' Explanations for Their Similarity Groupings*

Problem 6 (21)

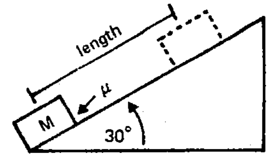


Expert 2: "*Conservation of Energy*"

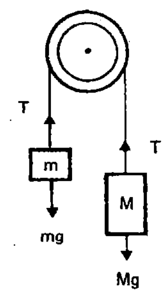
Expert 3: "*Work-Energy Theorem*. They are all straight-forward problems."

Expert 4: "These can be done from energy considerations. Either you should know the *Principle of Conservation of Energy*, or work is lost somewhere."

Problem 7 (35)



Problem 5 (39)



Expert 2: "These can be solved by *Newton's Second Law*"

Expert 3: "F = ma; *Newton's Second Law*"

Expert 4: "Largely use F = ma; *Newton's Second Law*"

Problem 12 (23)

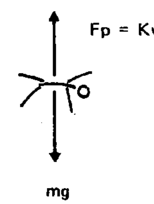


Figure 2. Diagrams depicted from pairs of problems categorized by experts as similar and samples of three experts' explanations for their similarity are provided. Problem numbers given represent chapter, followed by problem number from Halliday and Resnick (1974).

created. Tables 1 and 2 show the category descriptions (Column 1) used by more than one expert or novice. These category labels apply to all problems within each of their sorted piles.[1] Column 2 shows the number of subjects who used the category label. Column 3 shows the average size of the category among subjects who used it. And, Column 4 gives the total number of problems (out of 192, 24 problems for each of 8 subjects) according to category.

TABLE 1
Expert Categories

Category Labels	Number of Subjects Using Category Labels ($N_1 = 8$)	Average Size of Category ($N_2 = 24$)	Number of Problems Accounted for ($N_1 \times N_2$)
Second law	6	6.0	36
Energy principles (Conservation of Energy considerations, Work-Energy Theorem)†	6	5.5	33
*Momentum principles (Conservation of Momentum, Conservation of Linear Momentum, momentum considerations)†	6	5.0	30
*Angular motion (angular speed, rotational motion, rotational kinematics, rotational dynamics)†	6	3.0	18
Circular motion	5	1.6	8
*Center of mass (center of gravity)†	5	1.4	7
Statics	4	1.0	4
Conservation of Angular Momentum	2	1.5	3
*Work (work and kinetic energy, work and power)†	2	1.5	3
Linear kinematics (kinematics)†	2	1.5	3
Vectors	2	1.0	2
*Springs (spring and potential energy, spring and force)†	2	1.0	2

*Indicates the categories used by both novices and experts.
†When multiple descriptors across subjects were treated as equivalent, these are given in parentheses.

[1]For example, if a subject said of a problem group: "These all involve inclined planes, some with a frictional surface, some frictionless," the label "inclined planes" was counted since it applied to all problems in the set.

TABLE 2
Novice Categories

Category Labels	Number of Subjects Using Category Labels ($N_1 = 8$)	Average Size of Category ($N_2 = 24$)	Number of Problems Accounted for ($N_1 \times N_2$)
*Angular motion (angular velocity, angular momentum, angular quantities, angular speed)†	7	5.6	39
*Springs (spring equation, spring constant, spring force)†	6	2.8	17
Inclined planes (blocks on incline)†	4	3.8	15
Velocity and acceleration	2	5.5	11
Friction	2	5.0	10
Kinetic energy	4	2.0	8
*Center of mass (center of gravity)†	5	1.4	7
Cannot classify (do not know equations, do not go with anything else)†	4	1.8	7
Vertical motion	2	3.5	7
Pulleys	3	2.0	6
*Momentum principles (Conservation of Momentum)†	2	3.0	6
*Work (work, work plus Second Law, work and power)†	4	1.0	4
Free fall	2	1.0	2

*Indicates the categories used by both novices and experts.
†When multiple descriptors across subjects were treated as equivalent, these are given in parentheses.

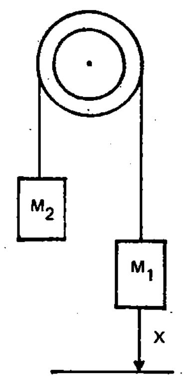
There are several things to note about these data which confirm our initial analyses of the four pairs of problems. First, there is little overlap between expert and novice categories. Only five of 20 distinct categories (marked with asterisks) are shared by the two groups. Second, if one considers the four predominant categories (the upper four in the tables in each subject group, ranked by total number of problems in each), the only overlap is in the category "angular motion." In particular, for these predominant classifications, the novices' descriptions are mostly objects and other surface characteristics of problems, whereas descriptions given by experts all involve laws of physics. Third, although both experts and novices classify a large number of problems (61 percent for the experts, 43 percent for the novices)[2] into four categories, there is a slight difference in the distribution of the problems across categories, which may suggest greater variability in novices' classification. Three major categories accounted for a sizable number (33 on the average) of experts' problems, whereas only one major category accounted for a large number (39) of novices' problems. This again suggests that experts are able to "see" the underlying similarities in a great number of problems, whereas the novices "see" a variety of problems that they consider to be dissimilar because the surface features are different.

Study Two: Sorting Problems with Surface Similarity

The objective of this study was to test our interpretations of Study One: experts categorize problems by laws of physics, and novices by surface features. A new

No. 11 (Force Problem)

A man of mass M_1 lowers himself to the ground from a height X by holding onto a rope passed over a massless frictionless pulley and attached to another block of mass M_2. The mass of the man is greater than the mass of the block. What is the tension on the rope?



No. 18 (Energy Problem)

A man of mass M_1 lowers himself to the ground from a height X by holding onto a rope passed over a massless frictionless pulley and attached to another block of mass M_2. The mass of the man is greater than the mass of the block. With what speed does the man hit the ground?

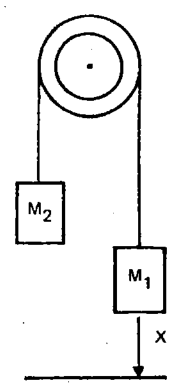


Figure 3. Examples of problem types.

[2]The percentages here do not correspond to those mentioned on page 5. Those were based on the largest sorting piles given by each subject, regardless of their contents or what they were labeled. Percentages here (Tables 1 and 2) are based on the sizes of specifically labeled categories when they were used by subjects.

set of 20 problems was constructed in which surface features were roughly crossed with applicable physics laws. Table 3 shows the problem numbers and the dimensions on which these problems were varied.[3] The left column indicates the major objects described in a problem. The three right headings are basic laws that can be used to solve problems. Figure 3 shows an example of a pair of problems that contain the same surface structure but different deep structure. In fact, they are identical, except for the question asked. Our prediction was that novices would group together problems that have the same surface structure, regardless of the deep structure, and experts would group together those problems with similar deep structures, regardless of the surface structure. Individuals of intermediate competence should exhibit some characteristics of each.

TABLE 3
Problem Categories

	Principles		
Surface Structure	*Forces*	*Energy*	*Momentum (Linear or Angular)*
Pulley with hanging blocks		20†	
	11	19†	
	14*	3*†	
Spring		7	
	18	16	1
			17+
		9	6+
Inclined plane	14*	3*†	
		5	
Rotational	15		2
			13
Single hanging block	12		
Block on block	8		
Collisions (Bullet-"block" or Block-block)			4
			6+
			10+

*Problems with more than one salient surface feature. Listed multiply by feature.
†Problems that could be solved using either of two principles, energy or force.
+Two-step problems, momentum plus energy.

[3]The problems were chosen from texts or constructed (to satisfy the a priori classification scheme) by Andrew Judkis, an assistant in the project who was a senior electrical engineering major with substantial experience in physics. Clearly, some problems could be solved using approaches based on either of two principles, Force and Energy, and in fact Judkis solved them both ways. In these cases, the problem is listed under the principle he judged to yield the simplest or most elegant solution but is marked with a cross. Also, some problems were two-step problems involving both momentum and energy. These are listed under the principle that seemed most important (in this case, momentum conservation) and are marked with a "+." These two-step problems are not designated explicitly as involving two principles. Some problems involve more than one potential physical configuration, e.g., "a pulley attached to an incline." These are marked with a single asterisk and listed multiply under alternative features.

The results confirm our previous interpretation. Table 4 shows the groupings and explanations of a novice who had completed one course in mechanics. This novice classification is based entirely on the surface structures of the problems. He collapsed problems across the physics laws, as was predicted. For example, the four problems in Group 2, 11 and 12 are force problems; 16 and 19, energy problems. The two problems in Group 4, classified by the novice as "Conservation of Energy," were problems purposely constructed as additional tests of "surface dependence" in novices; the novice identified them as energy problems only because they both have energy "cover stories" (i.e., they are stated in terms of energy), even though the major principle in each is Conservation of Momentum.

TABLE 4
Problem Categories and Explanations for Novice H. P.

Group 1:	2, 15	"Rotation"
Group 2:	11, 12, 16*, 19	"Always a block of some mass hanging down"
Group 3:	4, 10	"Velocity problems" (collisions)
Group 4:	13†, 17	"Conservation of Energy"
Group 5:	6, 7, 9, 18	"Spring"
Group 6:	3, 5, 14	"Inclined plane"
Groups 7, 8, 9 were singletons		

*Problem discrepant with our prior surface analysis as indicated in Table 3.
†Problems discrepant with our prior principles analysis as indicated in Table 3.

Table 5 shows the groupings of a physics graduate student. He classified the problems according to the three underlying physics laws specified a priori in Table 3. However, three of his classifications are discrepant with our analysis of the underlying principles. These discrepancies probably reflect deficiencies in his knowledge organization; the features in the problem statement cued the "wrong" category.

TABLE 5
Problem Categories and Explanations for Expert G. V.

Group 1:	3, 9, 2*, 17, 20, 5, 7, 19, 16	"Conservation of Energy"
Group 2:	13, 4, 10, 6, 15*, 1, 18*	"Conservation of Linear and Angular Momentum"
Group 3:	8, 12, 14, 11	"Statics problems or balance forces"

*Problems discrepant with our prior principles analysis.

That the graduate student's categorization was deficient is supported by Table 6 which shows the categories of a physics professor who sorted the problems after having spent considerable time thinking about how he would solve each problem in conjunction with a different task (reported in Study Four). Hence, this subject's categorization can serve as a validation for our prior analysis of problem types (Table 3). Only one problem, 9, is sorted according to a principle different from our choice.

TABLE 6
Problem Categories and Explanations for Expert V. V.

Group 1:	2, 13	"Conservation of Angular Momentum"
Group 2:	18	"Newton's Third Law"
Group 3:	1, 4	"Conservation of Linear Momentum"
Group 4:	19, 5, 20, 16, 7	"Conservation of Energy"
Group 5:	12, 15, 9*, 11, 8, 3, 14	"Application of equations of motion" (F = MA)
Group 6:	6, 10, 17	"Two-step problems: Conservation of Linear Momentum plus an energy calculation of some sort"

*Problem discrepant with our prior principles analysis.

What would an individual of intermediate competence do? Table 7 shows the groupings of an advanced novice (a fourth-year undergraduate physics major). His representations of the problems are characterized by the underlying principles in an interesting way. These principles are qualified and constrained by the surface components included in the problems. For example, instead of classifying all the force problems together (Groups 4, 6, and 7), as did the expert, he explicitly separated them according to surface entities of the problems. However, although he did not strictly group problems by physics laws, neither did he uniformly group them according to surface features. For instance, Groups 3 and 6 were separated even though they both involved springs. In addition, his groupings of principle were substantially discrepant with our prior analysis and that of the physics professor (Expert V.V. Table 6).

TABLE 7
Problem Categories and Explanations for Advanced Novice M. H.

Group 1:	*14, 20*	"Pulley"
Group 2:	1, *4, 6, 10*, 12*	"Conservation of Momentum" (collision)
Group 3:	*9*, 13*, *17, 18**	"Conservation of Energy" (springs)
Group 4:	*19, 11*	"Force problems which involve a massless pulley" (pulley)
Group 5:	2, 15*	"Conservation of Angular Momentum" (rotation)
Group 6:	*7*, 16**	"Force problems that involve springs" (spring)
Group 7:	*8, 5*, 3*	"Force problems" (inclined plane)

Italic numbers mean that these problems share a similar surface feature, which is indicated in the parentheses, if the feature is not explicitly stated by the subject.
*Problems discrepant with our prior principles analysis.