

Assessing Students' Misclassifications of Physics Concepts: An Ontological Basis for Conceptual Change

James D. Slotta, Michelene T. H. Chi, and Elana Joram
Learning Research and Development Center
University of Pittsburgh

Physics novices and experts solved conceptual physics problems involving light, heat, and electric current and then explained their answers. Novices were ninth-grade students with no background in physics; experts were two postgraduates in physics and two advanced physics graduate students. Problems were multiple choice, with one correct response and three alternative responses representing possible misconceptions. For each conceptual physics problem, an isomorphic *material-substance problem* was constructed by imagining a materialistic conception of the physics topic and creating the resulting version of the problem. In each physics problem, one of the incorrect choices corresponded to the correct choice in the isomorphic material-substance problem. The empirical question was whether novices would reason about the physics problem as if it were conceptually similar to the substance isomorph. This question was addressed by comparing subjects' responses in the problem pairs, as well as by examining their explanations concerning all problems. A content analysis of subjects' explanations revealed that physics novices were strongly inclined to conceptualize physics concepts as material substances, whereas expert protocols revealed distinctly nonmaterialistic representations. A theory of conceptual change involving ontologically distinct categories is substantiated by these findings.

For the past 2 decades, researchers have studied science misconceptions in a widespread effort to identify and characterize misconceptions of particular science concepts, as well as to experiment with different methods of modifying instruction so as to take these misconceptions into account. Pfund and Duit (1988) cataloged nearly 1,500 published studies that have identified a wide variety of students' misconceptions in science. Chi (1992) observed from this literature that some student misconceptions are easily removed in the course of instruction, whereas

others are characteristically robust. These "robust" misconceptions are typical of certain physics concepts and may be partly responsible for the difficulty experienced by both students and teachers in the physics classroom. If students are strongly attached to certain false conceptions of a physics concept, they may be unwilling to accept the conventional view of this concept offered by their instructor. For example, McCloskey (1983) observed that more than half the university engineering students he studied, all of whom had completed a semester of university physics, were plagued by misconceptions of the basic concept of force. In reviewing many such studies of physics misconceptions, Reiner, Slotta, Chi, and Resnick (1995) found that students often endow difficult concepts such as heat, light, or electric current with materialistic properties. A theory of conceptual change must take into account this disposition toward materialistic conceptions, as well as how such initial misconceptions can be corrected in the course of science instruction.

Research on science instruction has been largely unsuccessful in its goal of providing a theory of instruction (Resnick, 1983) that describes how students' misconceptions can be effectively identified and changed. To date, no instructional method has been successful in assuring that students entering the classroom will not finish the course with the same misconceptions they had on the first day of class. One reason for this lack of success is the corresponding lack of strong theories of conceptual change. Without at least one such theory, research is reduced to simply studying different instructional approaches, in contrast to predicting the effectiveness of a specific approach based on some theory of the underlying cognitive structures and processes. Another problem plaguing this research is the difficulty of assessing whether conceptual change has actually occurred. In a domain such as physics, it is entirely possible that a student may learn the "recipe" for correctly solving certain types of problems (e.g., Newton's-second-law problems) but retain misconceptions of the basic underlying concepts (e.g., force, velocity, weight). Thus, it is insufficient and perhaps inaccurate to claim that improved test scores on a quantitative problem-solving task can be taken as a measure of the extent to which student misconceptions have been removed.

A precise theoretical definition of *conceptual change* would be useful in determining a methodology for assessing whether and to what extent conceptual change has actually occurred. This article reviews one such definition, developed by Chi (1992; Chi, Slotta, & de Leeuw, 1994), together with a novel method of analysis that allows for assessment of conceptual change. The preliminary application of this method provides support for Chi's theory in terms of its ability to account for the relative difficulty of learning different classes of concepts.

CLASSIFICATION OF CONCEPTS

Many researchers (e.g., Medin & Smith, 1984; Rosch, 1978) have proposed that conceptual knowledge is organized hierarchically according to categories and that novel concepts are categorized by the learner according to similarity with

existing concepts. For example, a child might learn about a concept such as "bird" by forming a cluster of all the objects he or she has heard called *birds* and by eventually identifying the common attributes of all (or most) birds, such as "flies," "lays eggs," or "has a beak." A novel object might then be "categorized" as a bird if it appeared to have these relevant central attributes. Such a process of classification can result in a hierarchy of categories, defined by shared attributes. For example, all varieties of sparrow would belong to the *sparrow*¹ category, as well as to the higher level categories of *birds*, and then *animals*. Thus, if we categorize an unfamiliar object as a *bird* (based on its possession of the relevant attributes), it would then "inherit" many other attributes from higher level categories, such as "has blood," "breathes," and "dies," because these are all heritable attributes of the higher level category *animals*.

At some level of this hierarchy, categories can be said to be ontologically distinct, because the members of two different categories share almost no common attributes. Thus, although *sparrows* and *robins* are two categories whose members share many common features, *animals* and *minerals* have much less in common. Ontological attributes are those that apply to the most essential nature of a concept, such as whether it is living (e.g., a fish) or nonliving (e.g., a rock), an abstract idea (e.g., freedom), or a process (e.g., osmosis). Ontological differences between concepts are attended to by even very young children and are relied on as a source of inference (Carey, 1985; Gelman, 1978; Keil, 1979, 1989).

We have applied this notion of ontologically distinct classes of concepts to the issue of conceptual change in science learning by focusing on the distinction between two superordinate ontological categories: *material substances*² and *processes* (Chi, 1992; Chi et al., 1994). That is, students may classify science concepts according to these ontological categories and then rely on these classifications in subsequent learning as a source of inference or sense making. To the extent that students are mistaken in their ontological categorization of a particular concept, we propose that they exhibit characteristic misconceptions.

An ontological category is characterized by a set of *ontological attributes*, which are identified by verbal predicates that reflect ontological properties of the category's members. These are attributes that *may* but not necessarily *must* apply to concepts belonging to the category. For example, predicates that modify the ontological category of *material substances* include "heavy," "large," and "pourable," but any given concept belonging to this category, such as "canary," may not be heavy, large, or pourable. The critical aspect of ontological attributes is that they absolutely cannot apply to concepts in any other category. Thus, the sentence "The squirrel is broken" makes no sense to the listener, even when

¹Category names are italicized throughout the article.

²The category *material substances* includes physical objects of the *natural kind*, such as animals and rocks, as well as *artifacts*, such as machines and buildings. Thus, the *material substances* category subsumes other subcategories such as *animals* and *artifacts*. This is also true of the *process* category, which subsumes subcategories such as *interactions*, *events*, and *procedures*.

negated ("The squirrel is not broken"), because the attribute "is broken" comes from the *artifacts* category, which is ontologically distinct from the squirrel's actual category of *animals*. Contrast these ontological attributes with so-called "defining features." For example, a defining feature of canaries is their ability to fly, whereas this feature is not an ontological attribute of the *material substances* category to which canaries belong. Thus, ontological attributes are those that only category members may possess, as distinct from *characteristic* attributes (those that category members typically possess) and *defining* attributes (those that all category members must possess).³

We entertain the notion that ontological misclassification contributes to the difficulties experienced by students as they attempt to learn novel science concepts. Next, we discuss a theory of conceptual change that assumes that concepts are initially classified by the learner according to apparent ontological attributes. To the extent that such initial classifications are misconceived, they will interfere with subsequent instruction. We thus propose that some science concepts are more difficult to learn than others because they require a conceptual transition between ontologically distinct categories. Because this theory has been presented in detail elsewhere (Chi, 1992; Chi & Slotta, 1993; Chi et al., 1994), we provide only a brief discussion here, followed by some empirical results that support our hypotheses about the ontological commitments of physics experts and novices.

A THEORY OF CONCEPTUAL CHANGE

Why are certain concepts in science more difficult to learn than others, and by what underlying mechanisms does the learning actually proceed? The following simple theory of conceptual change asserts that concepts are associated with distinct ontological classes. On first encountering a science concept in the classroom or elsewhere, the learner classifies it according to its perceived ontological attributes. In further learning about the concept, the science student may be

³Ontological categories may be more or less distinct. For instance, the categories *solids* and *gases* appear to be ontologically distinct, but seem less divergent when compared with a category such as *ideas* or *processes*. That is, even though *solids* and *gases* appear to be ontologically distinct, many of their attributes apply to concepts in both categories (e.g., "has mass," "can be stored"). As discussed in Chi et al. (1994), this continuum of divergence can be accounted for by positing a hierarchical "tree" of categories that can be separated from one another vertically (as parent-child relations) or horizontally (as siblings). Vertically separated categories are ontologically compatible, because (by inheritance) a concept belongs to all categories of which it is a descendant. Horizontally separated categories can be more or less ontologically distinct, depending on the number of "parent" categories they have in common. For instance, the category *solids* is ontologically distinct from that of *liquids* but not to the same degree as it differs from the category *processes*. This is because the two categories, *solids* and *liquids*, presumably share some common ontology in the form of superordinate categories such as *material substances*. Chi et al. (1994) presented one possible hierarchical tree, although the theory does not depend on any particular version of the tree. This article requires no assumptions about hierarchies.

required to revise his or her categorization of the concept's ontology. If the concept's actual (i.e., textbook) ontology differs from the student's initial classification, the process of conceptual change may be very difficult. Other concepts may require only slight modification or refinement of the initial ontological classification. By appealing in this way to the initial (i.e., student) and target (i.e., textbook) ontologies for science concepts, our theory provides an account of why certain concepts are more resistant to instruction than others.

One example is seen when a novice initially classifies electric current as a *material substance*. In subsequent physics instruction, electric current is treated as a *process* of interaction, constrained by the potential difference between two points in a circuit. If the novice is committed to a conception of electric current as a kind of "juice" that flows through the wire, he or she will find it difficult to learn the ontologically distinct conception of electric current as a net migration of electric charge under disequilibrium conditions within the circuit.⁴ Thus, misconceptions in physics are accounted for (Chi, 1992) by an *incompatibility hypothesis*: A concept that is initially classified by a novice in the wrong ontological category will be incompatible with subsequent information about the concept, thus rendering it more difficult to learn.

Reiner et al. (1995) observed a prominent trend in the literature concerning misconceptions of heat, light, electricity, and force: In novices' conceptions, these concepts tend to be viewed either as material substances or as the property of some material substance. This suggests that a single ontological category—namely, *material substances*—is preferred above others by physics novices as a source of intuitive knowledge about these physics concepts.

Chi (1992; Chi et al., 1994) proposed that these physics concepts actually belong to the category of *constraint-based interactions*, meaning that they are best thought of as possessing ontological attributes of *processes* rather than of *substances*. A working definition of *constraint-based interactions* follows: a special type of process in which a defined system (e.g., an electric circuit) behaves according to the principled interaction of two or more constraints (e.g., the voltage at different points in the circuit). These principled interactions typically correspond to physical laws of nature, such as Ohm's law, Newton's second law, the laws of thermodynamics, or Maxwell's equations.

Another example of the ontological difference between the *substance* view and the *process* view of a physics concept can be seen in the history of scientific understanding of heat. Early chemists described heat according to the caloric theory, in which an object's temperature was proportional to the quantity of caloric contained in the object. This early view, which attributed to heat a *sub-*

⁴The *substance* model suggests that some material substance (e.g., "juice," "electricity," or electrons) is the current itself and flows from one end of a wire to the other. In contrast, the *process* model suggests that electric current is a process involving the entire circuit, including some material elements such as electrons. There is an important distinction between (a) material substances and (b) processes involving material substances. This distinction is discussed at greater length shortly.

stance ontology, was eventually replaced by a thermodynamic view, in which heat was seen as a *process* of molecular excitation. Thus, as scientists have gained understanding of concepts such as heat, light, and electric current, they have identified principles that reflect the concepts' *process* ontology.

Once classified according to ontology, a concept becomes associated with the category's pool of ontological attributes. This commitment to ontology gives rise to misconceptions, as novices draw on the associated attributes in their reasoning and explanations. Similarly, differences between novice and expert explanations will derive from corresponding differences in their ontological commitment for any given physics concept. This article explores these differences by making predictions according to Chi's (1992) theory of ontological classification. If novices have classified a concept as a *material substance*, their explanations should contain verbal predicates that correspond to the ontological attributes of that category. Likewise, to the extent that novices and experts have different classifications of a concept, they should rely on different verbal predicates in their explanations. The concepts of light, heat, and electric current are used throughout this article because they are notoriously difficult for physics students, who reliably prefer certain characteristic misconceptions, as observed in the classroom and in cognitive research.

ASSESSMENT OF ONTOLOGICAL DIFFERENCES

The preceding discussion has paved the way for two empirical questions. First, do physics novices conceptualize certain topics as if they were material substances? Although this was suggested by the literature survey of Reiner et al. (1995), no direct empirical test of the hypothesis has yet been performed, particularly with respect to the issue of whether and to what extent such materialistic commitments cut across concepts. To address this question, some method of assessment is needed, here called a *category test*, to determine whether a participant has conceptualized a topic according to a particular ontological category (e.g., *substances* or *processes*). Given such an assessment tool, we can then ask a second question concerning possible categorical differences between the conceptions of physics novices and those of physics experts.

The study reported here employs a category test that measures the language used by novices and experts as they solve and explain various simple physics problems. Participants were provided with fairly plausible problem situations, with responses available in multiple-choice form. Half the situations are called, collectively, the *physics-concept problems*, because they are concerned with one of three concepts: light, heat, or electrical current. The other half are referred to as the *material-substance isomorphs*, because each one is identical to a corresponding physics-concept problem, given that the concept is thought of as a material substance. For example, the material-substance isomorphs for electric-

current problems generally consisted of water flowing through pipes, corresponding to a material-substance version of "electricity" flowing through wires. We hypothesized that physics novices would respond to the physics-concept problems as if they were simply alternative versions of the respective material-substance isomorphs. Experts, on the other hand, were expected to perform very differently on the two sets of problems, because they would not classify the physics concepts as material substances. Detailed analyses of problem explanations were then performed to determine more precisely the nature of novice and expert ontological commitments.

METHOD

Participants

Novices were 9 ninth-grade students with no background in physics. Experts were 2 advanced graduate students and 2 postdoctoral fellows, all in physics. Both graduate students were engaged in dissertation research with at least 5 years of graduate training. Participants received course credit (in the case of novices) or cash payment (experts) for their participation.

Materials

Physics-concept problems were designed to provide the physics novice with relatively unfamiliar situations involving the concepts of light, heat, and electric current, so that he or she would need to reason about the situation to make a prediction about its outcome. Each of these physics-concept problems was paired with an isomorphic material-substance problem, which was designed to be consistent with a materialistic misconception of the corresponding physics-concept problem. For example, one of the physics-concept problems (concerning heat) required participants to predict which of two cups of coffee would be hotter 20 min after they were poured: the one in a styrofoam cup or the one in the ceramic mug, given that both were fitted with airtight lids (see Figure 1). The material-substance isomorph for this problem was created by imagining heat as a material substance that can penetrate through some barriers more easily than others.⁵ Thus, the isomorph for this problem consisted of two helium balloons made of different materials, one more porous than the other, and participants were asked to predict which would be more buoyant after several hours. Figure 1 provides the actual stimulus items, together with representative explanation protocols (for the physics-concept problem) from a novice and an expert subject. We predicted that novices would choose the same answer for both problems of any given isomorphic pair (e.g., choice *a* in both of the problems shown in Figure 1), demonstrating

⁵The misconception targeted by this problem was a conception of heat as a material substance that can escape or pass through the walls of its container.

Physics Concept: Heat (prob. 2a)

Two cups of hot coffee are poured, one into a styrofoam cup and one into a ceramic mug, and both cups are sealed with airtight lids. What will we find after leaving the two cups sit on a table top for twenty minutes?

- (N₁) a. The coffee in the ceramic mug is hotter than that in the styrofoam cup.
- (E) b. The coffee in the styrofoam cup is hotter than that in the ceramic mug.
- (N₂) c. Neither cup has hotter coffee than the other.
- d. Other

Material Substance Isomorph (Prob. 10a)

Two different balloons are filled with Helium gas, one made of an ordinary paper bag and one made of durable elastic rubber, and both are sealed tightly at the opening. What will we find after leaving the two balloons floating inside a closet for several hours?

- (E, N) a. The balloon made of rubber is more buoyant than the one made of paper.
- b. The balloon made of paper is more buoyant than the one made of rubber.
- c. Neither balloon is more buoyant than the other.
- d. Other

Novice Protocol (prob. 2a)

... A, the coffee in the ceramic mug is hotter than that in the styrofoam cup, because the heat in the styrofoam cup is gonna *escape* because a styrofoam cup is not totally sealed because there's like, styrofoam has little holes in it. So it, the heat is gonna *go out*, *escape*, in the holes. And the ceramic cup doesn't have, it's just totally sealed tight, and so I think it's gonna be hotter.

Expert Protocol (prob. 2a)

... B, the coffee in the styrofoam is hotter. Because my own experience is that the conductivity of the styrofoam cup is lower than that of the ceramic. Basically, for these types of materials in the solid form, the dominant source of energy loss is through conduction. Now, exactly, what is conduction? Conduction is *transfer* of heat from a hotter source to a cooler source, from, due to the *movement or motion of electrons*, and, in part, energy from the hot source to the cooler side of the material due to *excitation* of electrons which thereby collide with other nearby electrons and hence *impart energy* slowly from the hot side to the cool side.

FIGURE 1 One isomorphic pair of problems and two contrasting explanations of the physics-concept problem. *N* signifies novice misconception; *E* signifies expert choice.

a materialistic commitment in their conceptions of the physics concepts. Experts, however, were expected to correctly solve both the physics-concept problem (choice *b* in the first problem of Figure 1) and the material-substance problem (choice *a* in the second problem of Figure 1).

The various alternative responses for problems, such as the one shown in Figure 1, were derived from the different possible conceptions a student might have of the physics concept (heat, in this case). One choice was readily derived from the veridical "physics" conception of the problem (choice *b* in this example). Another (choice *a*) was obtained by imagining the concept as a material substance and predicting the consequences. A third choice (choice *c*) was used to logically balance the first two (e.g., choice *c*: "Neither cup has hotter coffee than the other"), thus providing the stimulus with a complete set of choices, as well as a control for guessing. Finally, choice *d*, "Other," was included in all our test items.

After the various alternative choices for the physics-topic problem had been designed, it was fairly straightforward to map these choices onto the material-substance isomorph. Development of the material-substance isomorphs was guided by the existing body of research concerning naive conceptions in physics (see Reiner et al., 1995). The choice of material substance for a particular physics concept thus corresponded to typical naive conceptions of that concept (e.g., flowing water for electric current). For each physics concept, 6 isomorphic pairs of problems were designed. Hence, there were 12 multiple-choice problems for each of the three concepts, creating a total of 36 problems in all.

Procedure

Before seeing any of the problems, participants received a set of prepared instructions, which were read aloud by the experimenter and included an example situation with a detailed explanation. Participants received an ordered stack of 36 problems and were asked to read along as the experimenter read aloud first the problem situation and then each alternative solution. Only then were they asked to choose one of the solutions. Problems were read aloud to control for possible differences in participants' reading ability. After they had selected a choice, participants were asked to explain their choice in terms of the problem's underlying causes and structures. That is, they were asked to explain not just their answer but also the entire problem. Feedback was given concerning the appropriate level of detail in explanations. All explanations were tape recorded, and participants experienced no apparent difficulty in understanding or executing the task. Although verbal probing by the experimenter was held to a minimum, participants were occasionally probed in the midst of their explanations when the explanation was unclear or when they had failed to offer a complete explanation of the situation.

Because the material-substance problems for a given physics concept tended to be very similar (e.g., for the electric-current problems, the isomorphs typically included water faucets, hoses, and water wheels), it was important to avoid

promoting any material-substance analogy for the physics concept (e.g., flowing water as an analogy for electric current) before all the physics problems for that concept had been completed. Therefore, participants were asked to solve and explain all 18 of the physics concept problems first, followed by all 18 of the material-substance isomorphs. Within each of these blocks, the problems from the three topics were counterbalanced for maximal separation of conceptually similar problems (i.e., two electricity problems were always separated by a heat and a light problem). Finally, despite the blocked presentation of the items, there were no obvious indications that novice participants perceived a discontinuity between the two sets of problems. That is, participants appeared to consider all 36 problems as a single homogeneous set.

RESULTS AND DISCUSSION

Multiple-Choice Data

Multiple-choice response data are shown in Figure 2. First, both novice and expert participants achieved nearly perfect performance on the material-substance problems. Although these problems were never intended to be difficult, this result is important, because novices and experts should be equally adept at understanding such problems (because both groups should have well-formed knowledge of the *material substance* ontology). Second, novices and experts differed markedly in their responses to the physics-concept problems, with experts achieving perfect performance (100% correct) and novices solving less than half the problems correctly (mean score of 45%, or 8.1 of 18 problems correct). Although, ideally, novices would solve none of the physics-concept problems correctly (they were predicted to always choose the material-substance misconception), they apparently

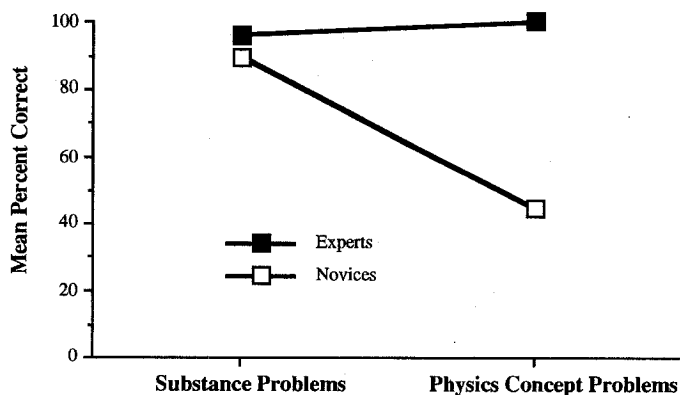


FIGURE 2 Multiple-choice data for experts and novices. Experts did not differ from novices in performance on material-substance problems but differed significantly in performance on physics-concept problems.

chose the correct response to some of the physics-concept problems that contained objects and situations familiar to them from their everyday experiences.

When novices did not choose the correct response to the physics-concept problems (as occurred on 55% or 9.9 of 18 trials), they did choose the solution corresponding to the material-substance conception of the problem more often (an average of 6.4 of those 9.9 problems) than the alternatives (choice *c*, which was another wrong answer, usually consistent with some materialistic conception of the problem, or choice *d*: "Other"). On the remaining fraction of physics problems (where they chose neither the correct response nor the material-substance conception), novices chose the third alternative, which was usually consistent with a materialistic conception of the problem but was ambiguous about the exact nature of participants' underlying conceptions. This ambiguity can be addressed by analyzing participants' explanation protocols, where they went into some detail about their understanding of the physics problems.

Perhaps because of familiarity with some problem situations, students can often choose the correct solution to physics problems without correct conceptual understanding (McCloskey, 1983; Reif & Allen, 1992). In the present study, novices chose the correct solution to the physics problem 45% of the time; in these cases, it is possible that they still relied on the *material substance* ontology as a source of conceptual understanding, a situation which should be reflected in the language they used to explain the physics-concept problems. For this reason, as well as to provide a second line of converging evidence, a detailed analysis of the explanation protocols was undertaken. The multiple-choice data have provided a useful demonstration of the kinds of misconceptions held by physics novices, as well as some qualitative support for our hypothesis that novices classify certain physics concepts as material substances. A detailed analysis of the explanation protocols across all 18 physics-concept problems (whether or not the correct solution was chosen) provides a second, more fine-grained test of this hypothesis.

Explanation Data

Explanation protocol analysis was guided by two basic questions. First, to what extent are novices' conceptual explanations based on a *material substance* conception of the relevant physics concept? Second, to what extent are experts' explanations based on a *constraint-based interaction* view of the relevant physics concept? To provide a glimpse of the explanation protocol data, as well as the method of coding described in this section, an Appendix has been included that presents several short excerpts from novice and expert explanations. Several issues associated with the interpretation of these data are also discussed at the end of the Appendix.

The first step in the analysis was to derive taxonomies of verbal predicates, which consist of words, phrases, or ideas whose presence in a spoken explanation

TABLE 1
 Taxonomy of *Substance* Predicates for Coding *Material Substance*
 Conceptions of Problems, With Examples From Protocols

<i>Substance Predicates</i>	<i>Examples</i>
Block	"keeps," "bounces off," "hits," "stops"
Contain	"holds in," "stores," "keeps in"
Move	"goes," "leaves," "comes," "flows through"
Rest	"stops," "stays," "sits"
Consume	"gets used up," "gets burned up," "burns out," "drains"
Absorb	"absorbs," "soaks up," "takes in"
Quantify	"some," "all," "most," "less," "none of," "lots," "little bit," "as much"
ColorAdd	"adds like colored paints," "red and blue make purple," "just like with paints"
Accumulate	"fills up," "builds up," "adds on," "keeps building"
Supply	"gives off," "provides," "comes from," "comes out of"
Equivalent amounts	"the same amount to all of the bulbs," "divides up equally"

is taken to reflect an underlying ontological attribute.⁶ Explanation protocols were coded for the presence of each taxonomy item by reading the protocol and deciding whether it contained verbal predicates that were equivalent in meaning to the taxonomy item. For the present analysis, two relevant taxonomies of verbal predicates were derived by enumerating ontological attributes of *material substances* (see Table 1) and *constraint-based interactions* (see Table 2). Tables 1 and 2 also provide some examples of their verbal predicate equivalents (words or phrases that were coded in the data as instances of a given taxonomy item).

In Table 1, the predicates "block" and "contain," for example, correspond to two attributes of material substances: They can block (e.g., another object) or be blocked (e.g., by a wall), and they can be contained. If, for example, in explaining a problem involving heat, a physics novice used expressions such as "the heat bounces off the walls of the cup" or "the heat can't get out of the cup," this would result in a coding of the "block" and "contain" taxonomy items from Table 1 and would thus be taken as support for the hypothesis that novices maintain a material-substance conception of heat.

Likewise, the predicates "transfer" and "interaction" in Table 2 correspond to attributes of processes: A process can be one in which transfer occurs (e.g., of energy or free electrons) or one of interaction (e.g., between electric and magnetic fields). Table 1 shows that it is often possible to express these predicates using

⁶The assumption that spoken predicates can be used as a measure of underlying cognitions is a very common one, but it is not without its criticism. Care must be exercised in the collection and interpretation of verbal data (for a complete discussion of the issues involved in analyzing verbal data, see Chi, in press; Ericsson & Simon, 1984), and any conclusions drawn from verbal data are subject to the same fundamental uncertainties associated with other correlative measures such as reaction times or eye movements. The assumptions particular to this article are addressed briefly in the Appendix.

TABLE 2
 Taxonomy of *Process* Predicates for Coding *Constraint-Based Interaction*
 Conceptions of Physics Problems, With Examples From Protocols

<i>Process Predicates</i>	<i>Examples</i>
Movement process	"... charged particle moving in an electric field," "the light is a traveling electromagnetic wave."
Transfer	"energy propagates through (the cup)," "... transfer from one to another."
Excitation	"a lot of phonon nodes to excite," "... need a lot of energy to excite them."
Interaction	"the interaction of the electric and magnetic fields," "the light energy is absorbed and transformed."
Equilibrium seeking	"The system finds its way into equilibrium."
Systemwide	"These are all in parallel," "... there's an electric field throughout the wire," "there's a field present throughout the wire," "all see the same potential."
Simultaneous	"They all see (the potential) at the exact same time."
Light as combined waves	"It would have red (spectral) lines and green lines in it."

words that are available to the novice. For example, the predicate "equilibrium seeking" could be expressed as "going until all parts are the same" or some such linguistic equivalent. Thus, novices would not need to know the actual word, *equilibrium*, to express this predicate, and often experts did not use the word itself.

Note that the individual members of the *constraint-based interaction (process)* taxonomy sometimes include material elements or qualities. In coding the explanations, it is important to be absolutely clear that our codes apply to a specific concept (e.g., heat, light, or electric current) and not simply to the language used in the explanations. Thus, we did not count every occurrence of the word *moves* as another case of the "move" predicate in Table 1. Rather, if we were coding an explanation of one of the heat problems, we counted only those occurrences of the word *moves* that directly applied to the concept of heat (e.g., "the *heat* moves/goes/comes").

This concept-specific coding enables us to preserve subtle but important distinctions such as the difference between the *substance* predicate "move" (Table 1) and the *process* predicate "movement process" (Table 2). For example, suppose we want to analyze a participant's ontological classification of the concept "wind" to determine whether the participant thinks of wind as a *substance* or a *process*. We would read the participant's explanation of some situation involving wind and code it according to the presence of any predicates from either Table 1 or Table 2. If the explanation included the sentence, "That sound you hear is just the wind moving through the canyon," this would result in a coding of the "move" predicate from Table 1, because the participant was talking about the wind moving. If, however, he or she had said, "The wind in the canyon is just the air moving," this would result in a coding of the "movement process" predicate

from Table 2, because he or she was talking about the wind as *moving air*. That is, the *air* was moving in this sentence (and, indeed, air is a *material substance*); the *wind* was described as the process of the air moving. Thus, even though the identical word, *moving*, appears in each of these sentences, each is coded differently. The predicate analysis used here is very sensitive to the conceptual content of the verbal data, because it attends not only to the presence of particular language elements but also to the context in which they are used.

After the taxonomies were established, each participant's explanation protocols were segmented into units corresponding in size to a single idea, which could range in size from short phrases to multiple sentences, depending on the destiny of the verbal data. Once the data had been segmented, they were coded according to the presence of any of the predicates defined in Tables 1 and 2. To simply count the resulting codes and draw conclusions at this point would have been erroneous because of possible confounds in the frequency data. Primarily, there was considerable interitem variability in terms of protocol length: Certain problems elicited more lengthy explanations than did others.⁷ Hence, the frequency measures were normalized to account for these differences. Normalization was accomplished in the following way for each participant: For each explanation, the raw frequencies for each predicate were tabulated (number of times a particular predicate was coded). These values were then divided by the total number of idea units in that particular explanation, so that the relative frequency of one predicate to another was preserved within a problem, whereas relative frequencies between problems (and participants) were adjusted to compensate for corresponding differences in the number of idea units used. The final measure thus reflects the actual importance of each individual predicate with respect to the explanation, as well as its relative importance to other predicates in the taxonomy. Thus, if a predicate appeared three times in an explanation that consisted of 10 idea units and another predicate appeared only once, the first predicate would still be considered 3 times more prominent in the analysis than the second.⁸

⁷Neither the material-substance problems nor the physics-concept problems elicited inherently longer explanations; a statistical test of the number of idea units included in explanations from the two blocks of problems showed no significant differences. Rather, certain problems in each block were more involved and required subjects to go into greater length in their explanations. Frequency data were normalized to offset this disproportionality.

⁸Two other methods of normalization were tried, both of which yielded very similar results. The first was to record the binary measure (1 or 0) of a predicate's occurrence. Thus, if a predicate were used in an explanation, it would be counted in full force compared with all other predicates that were used. This measure, although still yielding the same basic results, robbed the analysis of any sensitivity to differences between or patterns within the use of different predicates. A second method of normalization was to divide the frequency of a single predicate's occurrence by the total number of predicates used. This method resulted in nearly identical results to that employed in this article, because the number of idea units in an explanation was clearly proportional to the number of predicates coded. However, using the total number of predicates recorded as the denominator for normalization resulted in dependencies between the frequencies of individual predicates within an explanation, which we wanted to avoid.

The question of reliability in coding was addressed by having a second, independent rater code all 36 explanations from 2 participants (1 expert and 1 novice, with approximately 1,000 lines of transcribed protocol for each) for the presence of two *material substance* and two *constraint-based interaction* predicates.⁹ The rater was unfamiliar with the study and required approximately 1 hr of training, mostly regarding the *constraint-based interaction* predicates. Reliability was very high, with agreements of 90% for the *material substance* predicates and 84% for the *constraint-based interaction* predicates.

Predicate Frequency

Novices and experts were expected to reason similarly in explaining the material-substance problems. That is, we predicted that both groups would rely on the *material substance* (henceforth, *substance*) predicates and not on the *constraint-based interaction* (henceforth, *process*) predicates. In explaining the physics-concept problems, however, novices were expected to show no differences from their explanations of the material-substance problems, whereas experts were expected to abandon the *substance* predicates in favor of the *process* predicates. It is an important prediction that experts should not use many *process* predicates in explaining the material-substance problems, because they presumably do not categorize material substances as *constraint-based interactions*. Thus, we predict that novices and experts reason similarly in solving the material-substance problems but very differently in solving the physics-concept problems.

Figure 3 displays the frequency with which participants used *substance* predicates or *process* predicates in their explanations of the two types of problems (physics concept and material substance). Notice that the novices' use of verbal predicates does not appear to depend on the problem type. The two solid lines indicate that novices relied on the *substance* predicates (the filled squares) to the same extent in both types of problems, $F(1, 16) = 2.76, p > .1$, whereas they also ignored the *process* predicates (filled triangles) to the same extent, $F(1, 16) = 1.19, p > .1$.¹⁰ Differences between novice use of *process* and *substance* predicates were highly significant within both material-substance, $F(1, 16) = 309.1, p < .0001$, and physics-concept, $F(1, 16) = 537.3, p < .0001$, problems.

On the other hand, expert explanations appear to depend greatly on the type of problem. In Figure 3, the two dashed lines indicate that (a) experts relied on

⁹"Move" and "block" were the two *material substance* predicates; "process of excitation" and "transfer process" were the two *constraint-based interaction* predicates. A subset of the predicates was used because of the difficulty of searching in parallel for the entire taxonomy—a skill that required dozens of hours for the primary coder to build up. The training consisted mainly of providing made-up examples, as well as some physics instruction with regard to the *constraint-based interaction* predicates. It was most encouraging that, although the second rater was a physics novice, he was able to reliably code the expert explanation protocol for the presence of these two *constraint-based interaction* predicates.

¹⁰Novice frequency data were analyzed using a $9 \times 2 \times 18$ (Participants \times Problem Type \times Problem) repeated measures analysis of variance (ANOVA).

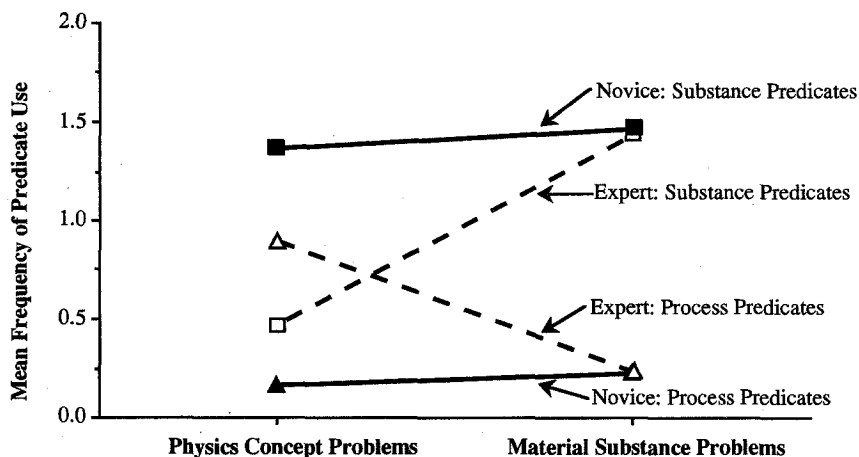


FIGURE 3 Use of *substance* and *process* predicates by 9 novices (solid lines) and 4 experts (dashed lines) for physics-concept and material-substance problems.

substance predicates (empty squares) to explain the material-substance problems but not the physics-concept problems, $F(1, 6) = 150.9, p < .0001$, and (b) experts relied on *process* predicates to explain the physics-concept problems but not the material-substance problems, $F(1, 6) = 29.2, p < .005$.¹¹ Differences between expert use of *process* and *substance* predicates were highly significant within both material-substance, $F(1, 16) = 198.1, p < .0001$, and physics-concept, $F(1, 16) = 19.3, p < .005$, problems.

The results displayed in Figure 3 support the hypothesis that novices classify certain physics concepts (light, heat, and electric current) according to the ontology of material substances. In contrast, experts appear to maintain a distinct ontology for these concepts, which is consistent with our hypothesized category of *constraint-based interactions*. Are these differences simply a result of experts having a more sophisticated language that contains many of our *process* predicates? As in any experimental paradigm that contrasts novice and expert behaviors, our results may have derived entirely from baseline differences (e.g., in age, education, life experience, IQ) between the two groups and not from the hypothesized differences in their respective conceptual frameworks. If novice and expert explanations differed solely because of language ability, however, the two groups should also differ in the language used to explain the material-substance problems. But experts and novices displayed very similar patterns of predicate use in their explanations of the material-substance problems: As shown in Figure 3, the two groups equally ignored the *process* predicates and equally preferred the *substance* predicates. Later, we show that this similarity extends further to

¹¹Expert frequency data were analyzed using a $4 \times 2 \times 18$ (Participants \times Problem Type \times Problem) repeated measures ANOVA.

their use of individual predicates in proportion to other predicates. As is always the case in expert–novice comparisons, the presence of a condition where experts and novices do not differ provides very helpful support against the claim that observed differences are simply a result of baseline effects.

The comparisons made in Figure 3 examine only the overall use of either *substance* or *process* predicates and say nothing about the use of individual predicates or, more important, the pattern of use among the predicates of a particular category. Pattern or distribution of predicate use is important because it provides an even finer-grained correlate of participants' underlying conceptual knowledge. In other words, the observation that novices use the same number of *substance* predicates in the physics-concept problems as they do in the material-substance problems says nothing about how they used these predicates. If they are indeed treating the physics-concept problems as being conceptually isomorphic to the material-substance problems, their explanations of the two types of problems should correspond. This correspondence can be explored by a finer-grained test of the *distribution* of predicate use within a category. In the next section, we discuss several such tests and how they relate to our hypotheses about ontological classifications.

Predicate Use Profiles

Grouped concepts. Bar charts, such as the ones displayed in Figure 4, are referred to as *predicate use profiles*, because they provide a means of visually comparing two distributions of predicate use. Figure 4a compares novice use of individual *substance* predicates in the physics-concept problems (solid bars) with their use within the material-substance problems (hatched bars). These two novice profiles appear to be identical, and a nonparametric test of correlation (Kendall Rank) between them was highly significant, $p < .005$ (see Table 3). Thus, novices drew on the same class of verbal predicates in explaining the two types of problems (Figure 3), and they did so with the same relative proportions, which adds further support for the hypothesis that they rely on a single ontological class of knowledge in order to construct these explanations.

Figure 4b illustrates the experts' use of *substance* predicates in the two types of problems. Clearly, both the frequency and the profile of experts' predicate use depend on the nature of the problem (i.e., material substance vs. physics concept). Thus, in explaining the physics-concept problems, experts did not make use of the *substance* category to the same degree (Figure 3) or in the same manner (Figure 4) as they did in explaining the material-substance problems. Moreover, it is highly relevant that the experts' *substance* predicate use profile for the material-substance problems (the shaded bars in Figure 4b) is very similar to that of the novices (Figure 4a): In explaining the material-substance problems, experts and novices used the *substance* predicates to the same degree (Figure 3) and in the same relative proportions (Figure 4). Figure 4 strengthens the conclusions implied by Figure 3 by showing that equal levels of predicate use within

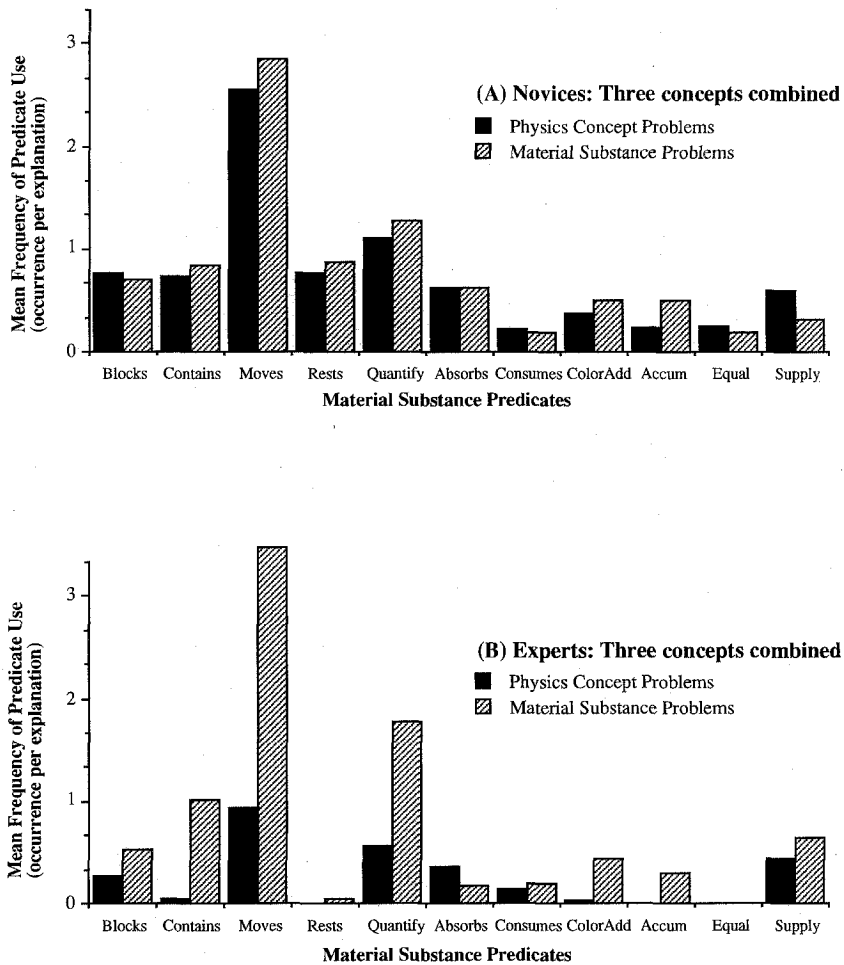


FIGURE 4 Distribution of *substance* predicate use according to individual predicates for (A) novices and (B) experts. Similarities in *pattern* of use reflect similarities in conceptualization.

a category (e.g., the *substance* predicates) are generally accompanied by equal patterns of predicate use as well. This finding provides further support for our assumption that experts do not differ from novices simply in the sophistication of their language, because the two groups made use of the same number and distribution of predicates in their explanations of the material-substance problems.

Individual concepts. Ideally, novices' predicate use profiles should correlate (between substance and physics problem explanations), even at the level of individual problems, not just when all 18 problems are pooled. That is, for a

single isomorphic pair of problems, the pattern of predicate use in the material-substance problem should resemble that of the physics-concept problem if the two problems are thought by the novice to be conceptually similar. Furthermore, these profiles should differ between different pairs of isomorphs, because the respective problems would draw on different ontological attributes of the *substance* category, depending on the specific topic of the different problems and their differing surface features. The sparsity of predicates in any single problem explanation and the small number of participants required by other tests prohibited any systematic analysis of predicate use in individual isomorphic pairs. It was still possible, however, to divide the 18 isomorphs into subsets according to the three physics concepts (6 isomorphic problem-pairs each for heat, light, and electricity) and to obtain enough coded predicates to perform a comparison across concepts. Distinct predicate-use profiles are predicted for the three physics concepts, which differ from one another but resemble the profiles of their associated material-substance isomorphs.

Figure 5 presents the novice explanation data from Figure 4a according to three separate profiles. Although the smaller number of data points for each profile is responsible for some increased variability, the effects are still visible: Each physics concept appears to have a distinct predicate use profile that presumably reflects the *substance* attributes most associated with that concept. Such differences are consistent with Chi's (1992) definition of ontological categories, which allows concepts from the same category to draw on different ontological attributes of that category. Moreover, variation within an ontological class is consistent with the existing literature on misconceptions. diSessa (1988, 1993) has argued that novice reasoning is influenced by the surface features of physics problems, which provide cues for the recall of phenomenologically derived knowledge fragments. The distributions presented in Figure 5 are consistent with diSessa's theory, because novices are cued by surface features of the problem, resulting in the use of different verbal predicates for problems with different physics concepts (e.g., heat vs. light vs. electric current). The theory thus provides an account of the *pattern of variability* observed in the predicate use profiles: namely, that a novice's explanations depend on the surface features of the problem, under the constraint of the ontological category according to which he or she has classified the relevant concepts.

The novices' predicate use profile for a given physics concept was similar only to the profile for the corresponding material-substance isomorphs. For instance, their use of predicates in reasoning about the electric-current problems was similar only to their predicate use in reasoning about the material-substance isomorphs for those problems and not to any other material-substance problems. The similarities suggested in Figure 5 are supported by the series of Kendall Rank Correlation tests reported in Table 3, which show that the predicate use profile for a particular topic correlated only with that of its own material-substance isomorphs (the diagonal cells). This pattern of correlations provides an even

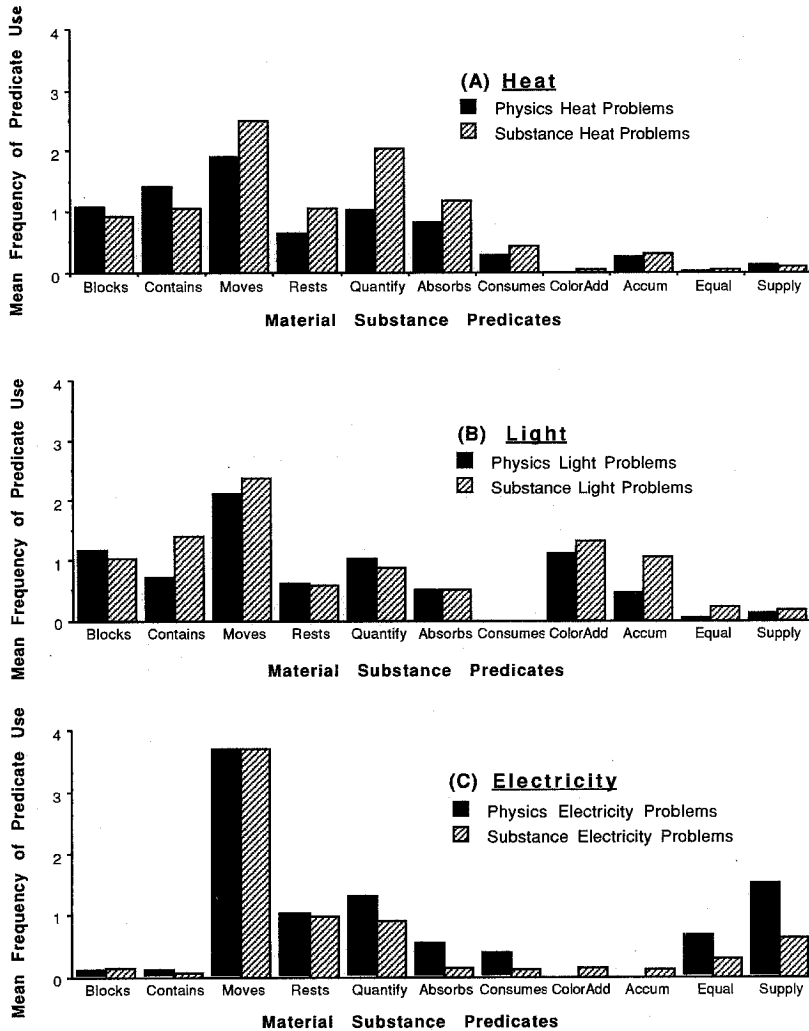


FIGURE 5 *Material substance* predicate use profiles of novices ($n = 9$) according to individual physics concepts. Distinct profiles for the three concepts are apparent.

finer-grained test of the hypothesis that novices classify these physics concepts according to the *material substance* ontology.¹²

¹²A similar correspondence between expert explanations for the individual concepts is not necessary. Although an analysis similar to that of Figure 5 would demonstrate this lack of correspondence, the view provided by Figure 4b is sufficient to show that experts' use of *substance* predicates was not correlated between the two types of problems. The obvious correspondence of novice explanations shown by Figure 4a led us to the finer-grained test of Figure 5. Similarly, the absence of such a correspondence in Figure 4b makes any finer-grained test uninteresting.

TABLE 3
Kendall Rank Correlations of 9 Novices' Use of *Substance* Predicates

<i>Substance Problems</i>	<i>Physics Problems</i>			
	<i>Electricity</i>	<i>Heat</i>	<i>Light</i>	<i>All Combined</i>
Electricity	.68 ($p = .005$)	.16 ($p = .484$)	.24 ($p = .312$)	
Heat	.37 ($p = .096$)	.81 ($p = .0005$)	.40 ($p = .084$)	
Light	-.13 ($p = .530$)	.38 ($p = .102$)	.67 ($p = .004$)	
All combined				.82 ($p = .0005$)

Note. Correlations of each concept with its own material-substance isomorphs (diagonal cells) are statistically significant, whereas correlations with other material-substance problems (off-diagonals) are generally not significant.

CONCLUSIONS

Novices and experts appear to draw on the same body of conceptual knowledge (the *material substance* category) in solving material-substance problems. Novices relied on this same knowledge when reasoning about the physics-concept problems, whereas experts displayed a categorically distinct body of conceptual knowledge in explaining their solutions to the physics-concept problems, in apparent agreement with our proposed ontology of *constraint-based interactions*. The novices' pattern of results was further supported over a fine-grained analysis that examined subjects' predicate use at the level of individual physics concepts.

The focus of this research has been to characterize the naive conceptions of physics novices and not to make any strong claims about experts' physics knowledge. That is, we sought to test particular claims about the source of robust misconceptions in science, including an apparent similarity across concepts that may derive from an underlying materialistic commitment (Reiner et al., 1995). In this article, we have provided a method of assessment that aims to confirm this conjecture by measuring patterns of language use in isomorphic pairs of problems. We developed pairs of problems (one physics concept, one material substance) that are only isomorphic if the physics-concept problem is misconceived as a material substance. Although novice responses to these pairs of multiple-choice problems suggested that such misconceptions were at work, a more probing analysis of the explanation protocols was required to provide a serious test of our hypotheses about naive ontological commitments. An analysis of expert protocols was included primarily to contrast with novice results rather than to make any strict claims about the "true" ontological nature of the physics concepts. Still, we are encouraged that the expert explanations revealed no sense of the *material substance* commitment demonstrated by novices but rather appeared to be consistent with our theoretical account of a commitment to the *process* ontology.

The theory of ontological categories affords a parsimonious account of the general trends observed in the literature on science misconceptions (Chi, 1992, 1993; Chi et al., 1994). The theory chiefly explains why some naive conceptions are robust (i.e., resistant to change) in so many respects:¹³ because novices form ontological commitments that prevent them from learning important aspects of the concepts. Our methodology of assessing predicate use, derived from the notion of ontological attributions, can be used as a measure of a student's ontological commitment and, therefore, as a criterion for conceptual change.

An even more elusive feature of naive conceptions is that, despite their overall robustness, they are highly susceptible to slight changes in context. For example, many researchers have noted that novices conceptualize heat as a material substance (Erickson, 1980; Tiberghien, 1979), but the precise nature of this conception appears to fluctuate according to the surface features of the particular experimental task. diSessa (1988, 1993) has argued that this variability is a defining feature of naive knowledge. Chi's theory can account for this context dependency in naive conceptions, because it suggests that novices draw on different ontological attributes, depending on the surface features as well as other conceptual requirements of the problem. Our results support this account, because the predicate use profiles are sensitive to the surface features of our experimental stimuli.

This work has important implications for developing students' conceptual knowledge in physics. If students enter the classroom with a firm commitment to substancelike conceptions, and if the theory of ontological classification is assumed, a particular course of action for instruction is implied. It will be important to avoid any possible attribute matches between the instructional materials and the *material substance* category. This can be done by adjusting the instructional language chosen, avoiding certain metaphors and analogies (e.g., the water analogy, which is very popular in modern instruction of electricity) and directly confronting the materialistic conception of a topic. It may also be important to provide some instances of the target category (e.g., *constraint-based interactions*), together with direct, high-level instruction about the definition of the category, so that the student is able to categorize the concept correctly.

ACKNOWLEDGMENTS

This work was supported by a grant from the Mellon Foundation. The opinions expressed do not necessarily reflect the position of the sponsoring agency, and no official endorsement should be inferred.

James Slotta is currently at the Department of Education, University of California at Berkeley. Elana Joram is currently at the Department of Psychology, University of California at Los Angeles.

¹³Chi (1992) argued that certain physics concepts are robust (i.e., resistant to change) irrespective of level of education, age, historical period, and learning context.

REFERENCES

- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Chi, M. T. H. (1992). Conceptual change within and across ontological categories: Implications for learning and discovery in sciences. In R. Giere (Ed.), *Cognitive models of science: Minnesota studies in the philosophy of science* (pp. 129–186). Minneapolis: University of Minnesota Press.
- Chi, M. T. H. (1993). Barriers to conceptual change in learning science concepts: A theoretical conjecture. In W. Kintsch (Ed.), *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society, Boulder, CO* (pp. 312–317). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Chi, M. T. H. (in press). Analyzing the content of verbal data to represent knowledge: A practical guide. *Journal of the Learning Sciences*.
- Chi, M. T. H., & Slotta, J. D. (1993). The ontological coherence of intuitive physics. *Cognition and Instruction, 10*, 249–260.
- Chi, M. T. H., Slotta, J. D., & de Leeuw, N. A. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27–43.
- diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufal (Eds.), *Constructivism in the computer age* (pp. 49–70). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10*, 105–225.
- Erickson, G. L. (1980). Children's viewpoints of heat: A second look. *Science Education, 64*, 323–336.
- Ericsson, K. A., & Simon, H. (1984). *Protocol analysis*. Cambridge, MA: MIT Press.
- Gelman, R. (1978). Cognitive development. *Annual Review of Psychology, 29*, 297–332.
- Keil, F. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.
- Keil, F. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299–324). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35*, 113–138.
- Pfundt, H., & Duit, R. (1988). *Bibliography: Students' alternative frameworks and science education* (2nd ed.). Keil, Germany: Institute for Science Education.
- Reif, F., & Allen, S. (1992). Cognition for interpreting scientific concepts: A study of acceleration. *Cognition and Instruction, 9*, 1–44.
- Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. (1995). *Materialism: An underlying commitment*. Manuscript submitted for publication.
- Resnick, L. B. (1983). Toward a theory of instruction. In S. G. Paris, G. M. Olson, & H. W. Stevenson (Eds.), *Learning and motivation in the classroom* (pp. 5–38). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Tiberghien, A. (1979). Modes and conditions of learning: An example. The learning of some aspects of the concept of heat. *Proceedings of an International Seminar on Cognitive Development and Research in Science and Mathematics* (pp. 35–55). Leeds, England: University of Leeds.

APPENDIX

This Appendix is provided to illustrate the verbal data, as well as to give a sense of our method of coding for predicate use. Two predicates have been selected from the *substance* taxonomy (“move” and “contain”), as well as from the *process* taxonomy (“movement process” and “excitation”). For each of these predicates,

we provide a brief definition and discussion, as well as a protocol excerpt from different participants for each of the three physics concepts (heat, light, and electric current). These protocol excerpts also serve to reflect the general qualitative nature of participants' explanations of the physics problems, including qualitative differences between novices and experts. Thus, all examples of *substance* predicate use are drawn from novice participants, because experts did not speak of these physics concepts as if they were *material substances*. Likewise, all examples of the *process* predicate use were taken from expert protocols, because novices rarely made use of these predicates. A short discussion of several issues involved in the interpretation of these data concludes this Appendix.

Substance Predicates

1. "*Move*." The predicate "move" is often applied to concepts that are *material substances*. For example, we might say that a train *moves* along the track. Because many other words imply movement, we would also consider sentences such as "The snowflake *falls* from the sky," "The man *comes* through the door," and "The ship *sails* across the ocean" to constitute valid occurrences of the "move" predicate. That is, any verb phrase that directly implies that the subject of the sentence (or of the sentential clause) is moving will be recorded as an instance of the "move" predicate. Because only *substances* can move, and *processes* cannot, we have suggested in this article that a sentence such as "The electric current *travels* along the wire" reflects a conceptualization of electric current as a *material substance*. Note that a statement such as "Electric current is when electrons travel along the wire" would not be recorded as an instance of the "move" predicate, because the concept of interest (electric current) is not being said to move in any way. Rather, the concept of electric current is being identified (the verb *is*) with a process of electrons moving along the wire. Thus, we would record this latter sentence as an instance of the *process* predicate "movement process," which is defined more precisely shortly.

The following protocol excerpts have been drawn from novice protocols to illustrate the types of statements that were coded in our analysis as instances of the "move" predicate for each of the three physics concepts. Particularly relevant key words within the explanation have been italicized to indicate the precise verbal data that we have interpreted as evidence of this predicate. Note that occasionally there is more than one italicized word or phrase in a protocol segment. Our "grain size" in this analysis was the "idea unit," which could range from a single phrase to several sentences in length, as determined by informational content. Even if more than one piece of evidence existed for a predicate within a single idea unit, the predicate was only coded once. Thus, in the first protocol segment to follow, which corresponds roughly to what we had segmented as an idea unit, there are two italicized indicators of the "move" predicate, but the predicate was only coded as occurring once here.

Heat

Novice 1: ... the coffee in the ceramic mug is hotter than that in the styrofoam cup, because the heat in the styrofoam cup is *gonna escape*, because it's not like, a styrofoam cup is not totally sealed because there's, like styrofoam has little holes in it, so it, the heat's *gonna go out*, escape in the holes so, and the ceramic cup doesn't have, it's just totally sealed tight ...

Light

Novice 5: ... when it's *traveling* from the, ah, flashlight, it's just *traveling straight forward* and then it would have to, ah, *go around the bend* ...

Electric Current

Novice 6: ... I think that there would be current *leaking out* of this one, 'cause if the current *was going* from the battery to the light bulb, and then through the part where it was cut, and *returning* to the battery, then I don't think there would be any from, past the cut to the battery ...

2. "*Contain.*" The predicate "contain" is another that can be applied only to *material substances* and not to *processes*. It refers to a situation in which some amount of a substance is held within a container. Once again, in our coding of the data we have recorded any verbal description of a concept that implies containment (e.g., "the heat is trapped inside") as an instance of this predicate.

Heat

Novice 9: So if you put something hot in there, it would *trap the heat* better than something ceramic, because ceramic doesn't have like air bubbles in it that can absorb the heat or the coldness ...

Light

Novice 2: ... a lot of the light wouldn't be going out, it would be staying in, so um, it would be kind of *trapping the light inside* so, um, it would just get, it would be brighter and then just stay that way ...

Electric Current

Novice 6: ... if you cut the circuit, there *would still be a little bit*, I think, of *current in the wire*.

Process Predicates

1. "*Movement process.*" As mentioned earlier, "movement process" is a predicate that attributes to the subject of a sentence (or sentential clause) the properties of a process in which some other object (not the subject) is moving. For instance, the sentence, "Wind is just the *movement of air*" would serve as an example of the "movement process" predicate if we were measuring predi-

cation of the concept "wind." In this case, the sentence describes wind as a process in which another object, air, is moving. In other words, the wind is not moving in this example; the air is. This would contrast with any sentence in which the wind is actually described as moving, such as "The wind *came down* the chimney," which would then be coded as an instance of the "move" predicate, described before. Although these distinctions are subtle, they provide important measures of the ontology attributed to these physics concepts. Physics experts have been shown in this study to prefer descriptions of electric current, heat, and light as *processes* in which some other object moves (i.e., electrons, molecules, and electromagnetic waves, respectively).

Heat

Expert 1: ... after the phase change to liquid state, then all of the energy the stove's pumping out is going into heating it up, which means it's *making the molecules of liquid vibrate more and more vigorously*, um, and it takes, once you shut off the stove, it's going to take longer for, for those, for the vibrational excitations to damp out, by transferring that again then to the air and to the stove and to the pan and everything.

Expert 3: ... the transfer of heat is not as efficient in glass, because ... metal has, ah, more of a crystal lattice structure so, we can have, ah, *phonons which just travel along* the lattice, moving up the lattice quicker. Um, phonons are really just lattice vibrations.

Light

Expert 2: ... light is a *form of energy that travels* faster than other forms of energy. ... It's a *physical process by which energy is transported* from one place to another.

Electric Current

Expert 4: ... when I say "current flow," what I really mean is that it is caused by the *movement of electrons*. Current is just generated by the *motion of electrons* which, in turn, is generated by a potential difference drawn across two different points in the wire.

2. "Excitation." The predicate of excitation refers to a process common in many physics systems in which some object within the system or the system as a whole is "excited" from one energy state to another. This energy could be in the form of atomic energy levels or of molecular kinetic energy.

Heat

Expert 4: ... heat is a form of energy ... that is being used to *agitate these atoms* even further, meaning that heat energy is being *transferred into* kinetic or rotational or to *even further linear kinetic energy, or translational kinetic energy* ...

Light

Expert 2: When I say "absorb the light," I mean some other atom, something, um, outside of the emitter, somehow *takes that packet of energy* and, ah, the atoms of that particular object, which is so called "absorbing" it, *takes the energy and uses it in some way in terms of exciting*, its, ah, electronic configuration, meaning the electrons that make up, that are part of the make up of that particular object are now *brought up to a higher level*, um, *energy state*.

Electric Current

Expert 3: ... The electrons see a potential difference across the light bulb, and they're *going to go to the lower energy position*; they're going to be forced to go.

There are several possible issues concerning the interpretation of these protocol data, most of which are related to the basic entanglement of knowledge and language. For instance, after reading these samples, one might conclude that the observed ontological differences between novices and experts are epiphenomenal, in the sense that they actually derive from obvious differences in physics expertise or linguistic ability. Thus, it might be argued that experts *speak* differently about these physics concepts because they possess the required physics knowledge, so that all we have done is record the obvious difference between those with physics knowledge and those without. The purpose of this study, however, was to *qualify* the nature of these differences according to a proposed measure of ontological commitment. A vital condition in the analysis (which is not illustrated in the preceding protocol excerpts) was the test of predicate use in the substance problems, where experts and novices displayed very similar patterns of predicate use. Thus, our observed differences were not strictly a consequence of baseline differences due to age, intelligence, or expertise. That is, experts spoke of "water moving through a pipe" just as novices spoke of "water moving through a pipe"; but, although novices spoke of "electricity moving through a wire," experts spoke of "an electric current being set up in a wire."

Another possible interpretation of these explanation protocols is that novices only *appear* to have material-substance conceptions but that this simply reflects their more limited linguistic abilities. That is, our ninth-grade novices may actually understand that "heat" is not simply a substance that builds up or leaks out of an object, but they are forced to *talk* about it as a substance because their limited linguistic skills are somehow biased toward materialistic descriptions. There are several reasons to believe that we have measured a true characteristic of naive conceptual knowledge and not just an epiphenomenon of ninth-grade language skills. First, this study employed a conceptual reasoning task in which a physics-concept problem (e.g., a heat problem) was isomorphic to a material-substance problem *only if* the physics concept were thought of as a *substance* (e.g., water). Clearly, this isomorphism is conceptually, not linguistically, rooted.

Novice responses to this task showed a clear preference for the material-substance view of the physics-concept problems (Figure 2), which suggests that they really were thinking of the physics-concept problem as if it were the material-substance isomorph. Second, our method of protocol analysis was designed to measure *patterns* of predicate use, which we interpreted as reflecting underlying conceptions. That distinct patterns of predicate use were observed in novice explanations for each of the three physics concepts and that the pattern for each concept correlated only with the pattern observed for that concept's material-substance isomorphs (Figure 5) suggest a correlation of the conceptual and not just the linguistic content of novices' explanations. Third, although the *process* predicates are linguistically more complicated than the *substance* predicates, this is not to say that they are linguistically beyond the reach of most ninth-grade students. The *process* predicates that we employed in our analyses could be expressed using fairly simple vocabulary and were often expressed in this way by experts. So, it was not the case that the *process* predicates were out of the novices' reach because they required big scientific words such as *excitation* or *propagation*, although such words were often employed by the experts. Finally, a simple qualitative reading of the novice explanations, which the excerpts presented here may allow, provides intuitive support for the claim that novices do think of these physics concepts as if they are material substances.

