Task Templates Based on Misconception Research

CSE Report 646

Jennifer G. Cromley and Robert J. Mislevy
National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)/
University of Maryland, College Park

December 2004

Center for the Study of Evaluation
National Center for Research on Evaluation,
Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 3.6 Study Group on Cognitive Validity, Strand 1 Cognitively Based Models and Assessment Design
Project Director: Robert J. Mislevy, CRESST/University of Maryland, College Park

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B960002, as administered by the Institute of Education Sciences, U.S. Department of Education.

TASK TEMPLATES BASED ON MISCONCEPTION RESEARCH[1]

Jennifer G. Cromley and Robert J. Mislevy
CRESST/University of Maryland, College Park

Abstract

Researchers spend much time and effort developing measures, including measures of students' conceptual knowledge. In an effort to make such assessments easier to design, the Principled Assessment Designs for Inquiry (PADI) project has developed a framework for designing tasks and to illustrate its use has "reverse engineered" several existing science assessments. This paper reports one such project, motivated by assessments that elicit students' qualitative explanations of situations that have been designed to provoke misconceptions and partial understandings. We describe four task-specific templates we created—three based on Hestenes, Wells, and Swackhamer's (1992) Force Concept Inventory and one based on Novick and Nussbaum's (1981) Test About Particles in a Gas (TAP). We then describe an overarching framework for these templates, another PADI object called a Design Pattern, based on Stewart's concept of "Model Using" (Stewart & Hafner, 1994). For each template, we describe a multivariate student model, a measurement model, and a task model. We conclude by suggesting how these templates and the design pattern could help researchers (and perhaps teachers) who wish to design new assessments in science domains where students are known to hold misconceptions.

1.0 Introduction

Creating tasks to assess underlying concepts and inquiry processes in science is not an easy thing to do. The National Science Foundation has funded the Principled Assessment Designs for Inquiry (PADI) project, under the Interagency Educational Research Initiative (IERI), to create a conceptual framework and supporting software to help people design inquiry assessments. Among the data structures PADI has developed to this end are *design patterns*, which lay out assessment arguments at a conceptual level; *task templates*, which are schemas for the operational elements of an assessment, and support the creation of families of related tasks; and *task specs*, which describe the elements of individual tasks in transportable formats (specifically, the IMS/QTI standards and extensions thereof).

One type of activity for the PADI project has been to take existing assessments used for science inquiry research, and to write templates and design specifications

[1] Thanks to Rick Elliott for preparing the manuscript.

for those assessments. We refer to this process as "reverse engineering," in that the templates and design specifications that are developed starting from existing tasks could be used to reproduce the same assessments, or to produce new or analogous questions in the same or a similar domain. This report presents the results of applying reverse engineering to two conceptual assessments in science: the Force Concept Inventory (FCI; Hestenes, Well, & Swackhamer, 1992) and the Test About Particles in a Gas (TAP; Novick & Nussbaum, 1981). Both assessments are based on research about student misconceptions, an area of cognitive psychology research on expert and novice performance.

Section 2 of the report provides a brief review of the theoretical basis for these assessments in the novice-expert and misconceptions paradigms in cognitive psychology. Section 3 then discusses the challenges that misconceptions research poses for assessment, and the benefits of making available to researchers the type of design patterns and templates that PADI is producing. Section 4 summarizes the structure of PADI task templates. Section 5 discusses the development of a design pattern and template for the Force Concept Inventory (FCI; Hestenes et al., 1992), a conceptual assessment of knowledge about Newtonian physics. The student model, evidence rules, statistical model, and task model for the template are discussed. Section 6 describes the process of adapting the template for the Test About Particles in a Gas (TAP; Novick & Nussbaum, 1981). Section 7 closes with a summary of the potential benefits to users of the design pattern and of the four templates that were developed.

2.0 Novice-Expert Research

One of the dominant strains of cognitive psychology research from the mid-1970s through the 1980s was the study of expert performance across numerous domains (for reviews, see Charness & Schultetus, 1999; Ericsson & Charness, 1997). The basic premise of this line of research was that if the characteristics of expert performance could be isolated and identified, then perhaps novices could be trained in those specific knowledge, skills, and attitudes, in order to move them closer to expert performance. Moving away from previous notions of expertise as general and inborn, cognitive psychologists conducted research that led them to see expertise as domain specific and acquired through extensive teaching and practice (Ericsson, 1996; Ericsson & Smith, 1991). Though early expert-novice research tended to consider only expert performance (e.g., in chess, de Groot, 1946/1978), later research often contrasted experts and novices (e.g., in physics, Larkin, McDermott, Simon, & Simon, 1980). Early research also focused on problem solving in domains with clearly delimited solutions and limited solution paths, called well-defined domains in the literature. Physics, chess, and medicine were particularly often studied, but other well-defined domains included the work of avionics technicians, waiters, and taxi drivers.

In physics, several researchers contrasted physics professors with typical undergraduate students. For example, Chi, Feltovich, and Glaser (1981) asked both physics professors and undergraduate students to sort various physics problems. Whereas the professors sorted the problems according to the physical laws that would be used to solve the problem (e.g., Newton's third law), novices tended to sort them according to physical features of the problem (e.g., pulley problems). In chess, Chase and Simon (1973) found that expert players were better able than novices to reconstruct the positions of chess pieces from memory, but only when the pieces were arrayed in actual game positions, not when they were randomly placed on the board. In medicine, Patel and Groen (1991) found developmental trends in medical expertise; whereas first-year medical students were not even aware that a written case contained irrelevant information, second-year students were distracted by that irrelevant information, and physicians recognized it as irrelevant.

In the course of developing a computer-based Intelligent Tutoring System (ITS) for military aircraft electronics technicians, Gitomer and colleagues found that expert technicians used a specific problem-solving strategy called they termed "space splitting" (Steinberg & Gitomer, 1996). If there was an electrical fault in a line, expert avionics technicians would pick a halfway point between the power source and the inoperable part (e.g., a wing flap or aileron), and would test the circuit on both sides of the line. By continuing this process, they could quickly identify the malfunctioning electrical component. Novice technicians, by contrast, would use an inefficient process of checking each component in turn, one at a time. Ericsson and Polson (1988) found that the expert waiter whom they studied had developed detailed heuristics (e.g., males with certain builds were likely to order certain types of steaks) and mnemonics for remembering diners' orders. Expert taxi drivers studied by Chase (1982) used mental imagery to determine the quickest route from the current location to their destination.

These studies in well-defined domains established that expertise takes many years of guided practice to develop, that experts have a larger base of declarative knowledge than do novices, that their knowledge is better integrated and organized according to key principles in the domain, that experts know more domain-specific strategies, that their use of these strategies is automated, and that experts have more conditional knowledge about strategies (that is, they know in what situations to enact a particular strategy).

A later body of expert-novice research applied these findings to domains that do not have a single solution or finite set of solution strategies (called ill-defined domains), such as reading (e.g., reading law cases, Lundeberg, 1987; reading poetry, Peskin, 1998), history (e.g., Wineburg, 1991), teaching (e.g., Sabers, Cushing & Berliner, 1991), and writing (e.g., Breuleux, 1991), with similar results.

Typically, in novice-expert studies researchers collect some sort of verbal report from participants, such as a think-aloud protocol (participants verbalize everything they are doing while performing a task) or a retrospective protocol (participants report after the fact what they think they were doing), or conduct an interview. A few methodologists have suggested that expertise researchers should simultaneously collect other process data such as recording computer keystrokes, recording what participants are looking at (using eye-tracking devices), or reaction times (Ericsson & Smith, 1991; Magliano & Graesser, 1991). Researchers have also used other methods such as the sorting task described above for Chi et al.'s (1981) physics expertise study.

One common finding across many expert-novice studies, especially in the sciences, is specific mistaken ideas that novices have about particular domains. In one highly publicized example, 88% of graduating Harvard seniors surveyed believed that the seasons were caused by Earth's elliptical orbit around the sun, rather than by the tilt of the Earth on its axis (Gardner, 1999). Termed misconceptions, these are ideas derived from daily experience that students bring to their learning experiences, and which contradict scientific understandings.

Some other examples of misconceptions include these: in biology, exercise makes breathing faster but shallower (e.g., Michael, 1998); in subtraction, subtract the smaller number from the larger, regardless of which number is being subtracted from, for example, 725 – 569 = 244, since 7 - 5 = 2; 6 – 4 = 4; and 9 – 5 = 4 (e.g., Brown & vanLehn, 1980); in history, textbooks are always correct and we can know "what happened" (e.g., Rouet, Britt, Mason, & Perfetti, 1996); in evolution, animal behavior changes offspring biology (a Lamarckian belief; e.g., Bishop & Anderson, 1990); and in Earth science, the Earth is a globe with a flat top (e.g., Vosnaidou & Brewer, 1992). Research about misconceptions, thus, has been a wide-ranging and productive field for cognitive science research, and one which has been frequently used in science inquiry research.

3.0 The Challenge for Assessment

Assessing misconceptions poses great challenges for test designers. To begin with, misconceptions are defined as misunderstandings that are not detected by traditional assessments. Physics students, for example, may be able to use algorithms to calculate answers to problems about which they have no conceptual understanding (e.g., Clement, 1989; Hunt & Minstrell, 1994). Assessments designed to detect misconceptions, therefore, usually eschew calculation problems, on the grounds that these can be solved without having any conceptual understanding of the problem. In addition, because the purpose of assessing misconceptions may be more diagnostic than normative, researchers (and teachers) may want to know about which specific misconceptions students have (see Frederiksen & White, 1988). This type of use for assessment therefore requires different types of measurement models (and perhaps also student models) than does an end-of-semester physics examination used for assigning grades.

Researchers who are studying science inquiry may also be particularly interested in measuring misconceptions, since inquiry activities (more so than didactic or cookbook lab approaches) may reveal students' misconceptions (e.g., Dalton, Morocco, Tivnan, & Mead, 1997; White & Frederiksen, 1998). However, developing psychometrically sound measurement instruments for science misconceptions is a difficult and time-consuming task, and one at which test developers are not always successful (see Cornely-Moss, 1995). Researchers who study science inquiry may be spending much time, energy, and resources "reinventing the wheel" as they develop measures for individual projects. In addition, researchers rarely have access to the kind of sophisticated statistics and technology that make both student models and statistical models more accurate and efficient. The goal of the Principled Assessment Designs for Inquiry (PADI) project is to address these needs by creating structures including design patterns and templates to help assessors organize their thinking about science learning into the shape of assessment arguments and tasks, and by illustrating their use in a variety of contexts. In this technical report, we present a series of templates and one design pattern that encompass two science domains: Newtonian physics and gas laws.

4.0 PADI Design Patterns and Task Templates

The PADI structures called design patterns and task templates build on the "evidence centered" assessment design (ECD) models of Mislevy, Steinberg, and Almond (2003). A good starting point is a quote from Messick (1994):

> A construct-centered approach [to assessment design] would begin by asking what complex of knowledge, skills, or other attribute should be assessed, presumably because they are tied to explicit or implicit objectives of instruction or are otherwise valued by society. Next, what behaviors or performances should reveal those constructs, and what tasks or situations should elicit those behaviors? Thus, the nature of the construct guides the selection or construction of relevant tasks as well as the rational development of construct-based scoring criteria and rubrics. (p. 17)

A PADI design pattern lays out, at a conceptual level, coherent sets of possibilities for the elements of the assessment argument outlined in the Messick (1994) quotation, organized around some aspect of scientific inquiry or conceptual knowledge. For example, there are the targeted knowledge, skills or other attributes (targeted KSAs); things that students might say, do, or make that provide evidence of these KSAs (potential observations); and characteristic features of task situations in which these observations might be made. The design pattern structure is described and illustrated in detail in PADI Technical Report 1, *Design Patterns for Assessing Science Inquiry* (Mislevy, Chudowsky, et al., 2003). This brief description, however, should suffice to understand the example in Section 6.

More attention is focused in this presentation on task templates. Again the reader is referred to PADI Technical Report 2, *An Introduction to PADI Task Templates* (Riconscente, Mislevy, & Hamel, 2004). Task templates are organized around three basic models in the Mislevy, Steinberg, and Almond (2003) assessment framework, namely, the Student Model, the Evidence Model, and the Task Model:

- The Student Model contains variables that correspond to knowledge, skills, and abilities of an examinee about which inferences will be made--decisions about selection, placement, certification, instruction, task selection, and so on.

- The Evidence Model is a set of instructions for interpreting the response (Work Product) to a specific task. The Evidence Model contains two parts. The first is a series of Evaluation Procedures that describe how to identify and evaluate essential features of the Work Product. The second is a Measurement Model that tells how the belief about the student model variables for a given student should be updated in light of the observed features of that student's responses.

- The Task Model is a generic description of a family of tasks. A Task Model contains (1) a list of variables that are used to describe key features of the tasks, such as their content, difficulty, and conditions under which they are presented; (2) a collection of Presentation Material Specifications that describe the structure and format of material that will be presented to the participant as directions, stimulus, prompt, or instruction; and (3) a collection of Work Product Specifications that describe the structure and format of material that the task will be evaluated.

The graphic shown as Figure 1 shows the constituent objects in a PADI template, which together encompass the student model, evidence model, and task model described above. The interested reader is referred to Riconscente et al. (2004) for detailed explanations of the structure of templates. This diagram, however, will help in understanding the examples of templates presented in the appendices, by indicating the hierarchical structure of the pieces that comprise a template.

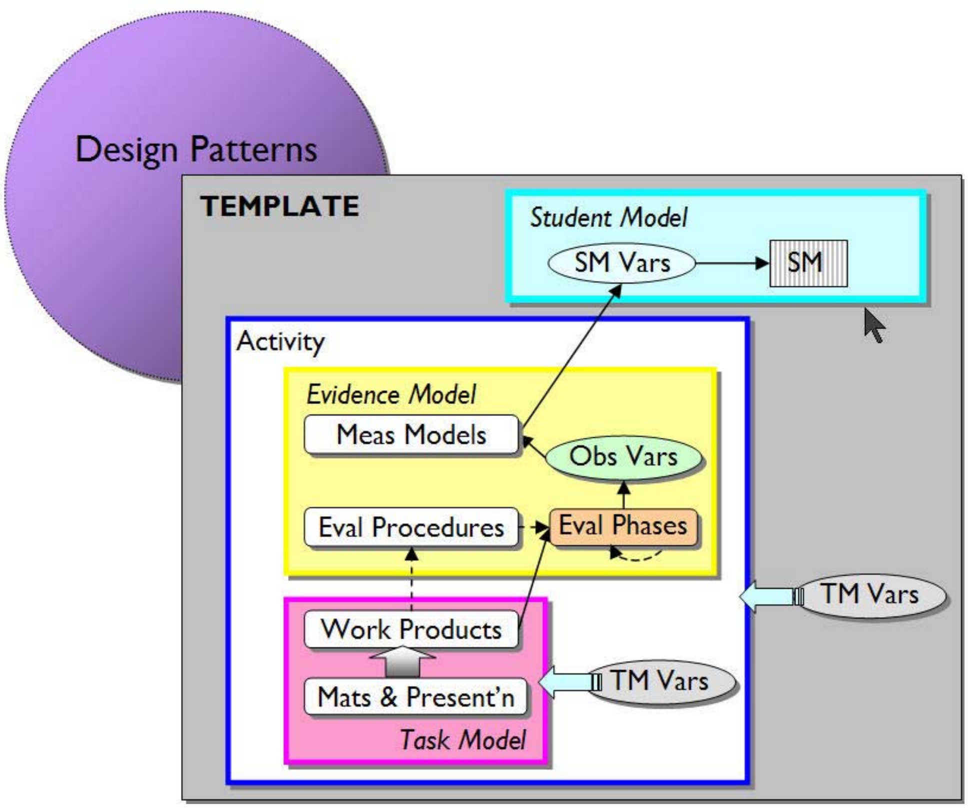


Figure 1. The hierarchical structure of PADI templates (from Riconscente et al., 2004).

5.0 Illustrations From Newtonian Physics

One domain in which much misconceptions work has been done is Newtonian physics. Several studies have shown that students can solve typical classroom quantitative problems in Newtonian physics, while failing to understand basic Newtonian principles (see, e.g., Clement, 1983; diSessa, 1993; White & Frederiksen, 1998). Newtonian motion problems can be divided into problems that consider an object moving in a horizontal plane (e.g., a hockey puck moving across ice), other problems that may concern only gravity acting on an object—either a still object or a moving object, and a third class of problems that concern horizontal motion and the vertical force of gravity acting simultaneously. A typical undergraduate physics assessment question asks whether students can correctly compute the time, distance, acceleration, or force needed to move an object:

> A projectile is fired horizontally from a flare gun located 45.0 m above the ground. The projectile's speed as it leaves the gun is 250 m/s.
>
> a) How long does the projectile remain in the air?
> b) What horizontal distance does the projectile travel before striking the ground?
> c) What is its speed as it strikes the ground?
> d) If the projectile were simply dropped from a height of 45.0 m, instead of fired horizontally from that height, how much time would it take to reach the ground? How does this compare with your answer to part (a)?
>
> (From http://ist-socrates.berkeley.edu:7521/projects/IPPS/Ch4/Prob6/Q.html, question 4-6)

Some common misconceptions include "impetus," the pre-Newtonian notion that as long as a body is in motion, a force must be acting on it (McCloskey, 1983). Students may be able to correctly calculate the amount of force required to set a body in motion—for example, to move a 10 kg box from rest—but they might believe that force is required to keep the box moving. In actuality, once the box has accelerated to its final velocity (assuming a frictionless world) no force is required to keep it in motion. A related misconception is that heavier objects exert more force (it does, of course, require more force to start a heavier object in motion; diSessa, 1993). A complementary misconception is that objects at rest must have no forces on them (Clement, 1983). In actuality, when an object such as book sits on a table, gravity (a force) pulls down on the book, and the table pushes back up with exactly the same amount of force against gravity.

Another misconception is the belief that heavier objects fall faster than lighter ones (the belief tested in Galileo's perhaps apocryphal experiment at the leaning tower of Pisa). Yet another misconception might be labeled the "Wile E. Coyote" misconception—the notion that "Any body suspended in space will remain in space until made aware of its situation" (from http://funnies.paco.to/cartoon.html). A student could likewise be able to calculate the horizontal distance that an object travels, but misunderstand that objects that move both horizontally and vertically always move in a parabola.

5.1 The Force Concept Inventory

In 1992, David Hestenes and colleagues published a multiple-choice measure of students' conceptual knowledge of Newtonian physics called the Force Concept Inventory or FCI (Hestenes et al., 1992). The FCI has been widely used as an undergraduate physics pre- and posttest to measure whether students truly understand motion or whether they can simply do calculations but lack a fundamental understanding. The FCI consists of 30 multiple-choice Newtonian physics problems that do not require any calculation, but are designed to tap students' understanding of various aspects of Newtonian mechanics and circular motion. The measure was constructed by reviewing prior research on correct Newtonian conceptions and specific student misunderstandings. In the original publication, Hestenes et al. list 23 specific Newtonian force concepts (Hestenes et al., 1992, Table I) that were used to construct correct options for the FCI and 30 specific misconceptions (Hestenes et al., 1992, Table II) that were used to construct incorrect options.

The following sections describe elements necessary for task templates that describe FCI-type tasks—that is, reverse-engineered templates that might be thought of as more general structures from which the actual tasks could have been derived. More substantive descriptions follow; "designer view" screen shots of actual PADI templates appear in the Appendices.

5.2 A Student Model

Hestenes et al. (1992) originally conceived of a univariate student model—students are Newtonian, either to a greater or lesser extent (i.e., they answer more or fewer FCI questions correctly). Bao and Redish (2001) suggested an alternative interpretation—that students can be in multiple belief states (e.g., Newtonian; Galilean, or "impetus" reasoning; and nonscientific, including Aristotelian reasoning) simultaneously, each with a certain propensity. This approach is similar to those of other multivariate models used in developmental and cognitive psychology. For example, Robert Siegler (1976) analyzed children's developing ability to solve problems in which different weights are put on either side of a balance beam. Children at the same stage of development may give different answers to the same balance beam problem, and children at different stages may give similar answers. However, overall, children in the same stage show a similar distribution of answers that are characteristic of their developmental stage. Thus, with the FCI we could imagine a student who answers like a Newtonian 70% of the time, like a Galilean 20% of the time, and in a nonscientific manner 10% of the time. Which kind of answer a student would give depends not only on the student's belief states, but also on the features of the task at hand.

In the PADI Design System, this student model has been named the "FCI-ish student model." Students have propensities to respond to tasks in a certain class (e.g., FCI and FMCE items) in terms of three specific conception models. These conceptions are Newtonian, Galilean, and nonscientific. Formally, each student i is characterized by a vector of three real numbers, (θ_{i1}, θ_{i2}, θ_{i3}), where higher numbers indicate a greater propensity toward a given response class—in this case, Newtonian, Galilean, and nonscientific. Statistically, the model can be made identified by implicitly fixing the sum of each examinee's three parameters at zero, or fixing the first to zero. The latter approach is taken in the example illustrated here, so there are only two student model variables to be estimated, namely for the Galilean and nonscientific propensities.

5.3 Evidence Model

Evaluation procedures and evidence rules. The FCI was originally a multiple-choice measure; evidence about the student model consisted of which multiple-choice option was chosen. With the multivariate student model approach we have taken in PADI, these evidence rules are expanded. Specifically, each multiple-choice option has been mapped to known Newtonian conceptions (Hestenes et al., 1992, Table I) and non-Newtonian misconceptions (which are further subdivided into the Galilean and nonscientific misconceptions contained in Hestenes et al., 1992, Table II). While we have chosen to analyze the FCI in its original multiple-choice format, it would be easy to adapt it for open-ended responses which would be coded. These codes would then be mapped to specific conceptions. We will see later that just this approach was taken by the designers of an analogous measure, the Test About Particles in a Gas (TAP; Novick & Nussbaum, 1981).

Measurement model. The measurement model for the FCI was originally a simple additive model. A multivariate statistical model for the FCI was recently investigated by Kevin Huang (2003) in dissertation research. Huang used an Andersen/Rasch (A/R) multivariate model (Andersen, 1973), in which students are seen as being in several belief states, each with a certain propensity, and specific FCI questions are seen as provoking certain belief states, each with a certain tendency to provoke responses of the different types.

The A/R model takes the following form for a situation in which there are m response types, and student and task parameters that correspond to them. Let X_{ij}; $i = 1, \ldots, n$, $j = 1, \ldots, k$ be independent observable random variables (i is the index for examinees, j is the index for items), where X_{ij} can be any integer between 1 and m. The probability that the response X_{ij} that student i produces for Task j is of Type m is given as

$$P(X_{ij} = p) = \exp(\theta_{ip} + \beta_{jp}) / \sum_{p=1}^{m} \exp(\theta_{ip} + \beta_{jp}) \,,$$

where:

> p is an integer between 1 and m, indicating response class;
> θ_{ip} is the pth element in the person i's vector-valued parameter; and
> β_{jp} is the pth element in the item j's vector-valued parameter.

Note that there are m probabilities for each examinee on a given item, representing the probability of choosing any particular choice for that person on that item.

The Andersen/Rasch model can be written as a special case of the MRCLM model used by the PADI project (Adams, Wilson, & Wu, 1997). For more information, see Huang (2003).

5.4 Task Models

In the PADI Design System we have created four task templates. All have the same student and evidence model structures, but differ as to the task models. Three task models are for particular types of FCI problems (Hestenes et al., 1992), and one is for the TAP (Novick & Nussbaum, 1981). In these task models, we do not create new items or content that goes beyond the work done by the authors of the measures. Rather, we write a more general set of task template structures (similar to item specifications) that can enable researchers—and perhaps teachers—to create analogous measures in their own domains of interest.

The three templates for FCI-type problems correspond to (a) problems involving only gravity, but no horizontal motion, which we refer to as G Problems; (b) problems involving only horizontal motion, but which do not tap students' knowledge of gravity, which we refer to as H Problems; and (c) problems which tap students' knowledge of both horizontal motion and gravity, which we call H & G Problems.

G Problems. A typical G Problem from the FCI is shown in Figure 2. The template entitled *Force Problems—Gravity Only* refers to the direction of motion of the object (up, down, or no motion), the duration of motion, friction, the identity of the vertical force (e.g., gravity, the force of an elevator cable pulling up), whether or not there is an illustration used in the problem, the mass and identity of the object being moved, the speed at which the object is being moved, and the time period of interest (e.g., how long does it take an object to fall). (See Appendix A for a screen shot of the Template and Task Model Variables from the *Force Problems—Gravity* template.) For example, the problem shown in Figure 2 involves the continuous upward motion of an elevator at an unspecified speed, no friction, the upward force of the elevator cable, the downward force of gravity, and an illustration.

Some of these task model variables are important in describing incidental features of the task situation, which may be varied to provide a range of tasks that are different on the surface but similar as to the thinking they tend to evoke. The identity of the object is such a variable. Other task model variables are important because they characterize features that are linked to common misconceptions. Direction of motion—whether an object is at rest, moving up, or moving down—is an example of this latter type. Many students believe that when an object is at rest, there are no forces acting on it.

G problems in the FCI ask about the forces acting on an elevator being pulled up a shaft, a steel ball that has been tossed straight up, a stone dropped from the roof of a building, a tennis ball at an instant after it has been hit, a metal ball at an instant inside a tube, and an office chair at rest on a floor. This particular template helps a researcher to create more FCI-like G Problems. The researcher creates a

17. An elevator is being lifted up an elevator shaft at a constant speed by a steel cable as shown in the figure below. All frictional effects are negligible. In this situation, forces on the elevator are such that:

(A) the upward force by the cable is greater than the downward force of gravity.

(B) the upward force by the cable is equal to the downward force of gravity.

(C) the upward force by the cable is smaller than the downward force of gravity.

(D) the upward force by the cable is greater than the sum of the downward force of gravity and a downward force due to the air.

(E) none of the above. (The elevator goes up because the cable is being shortened, not because an upward force is exerted on the elevator by the cable).

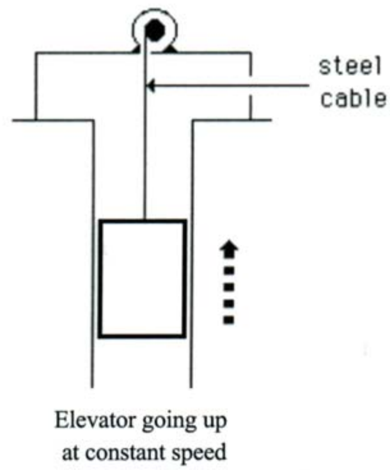
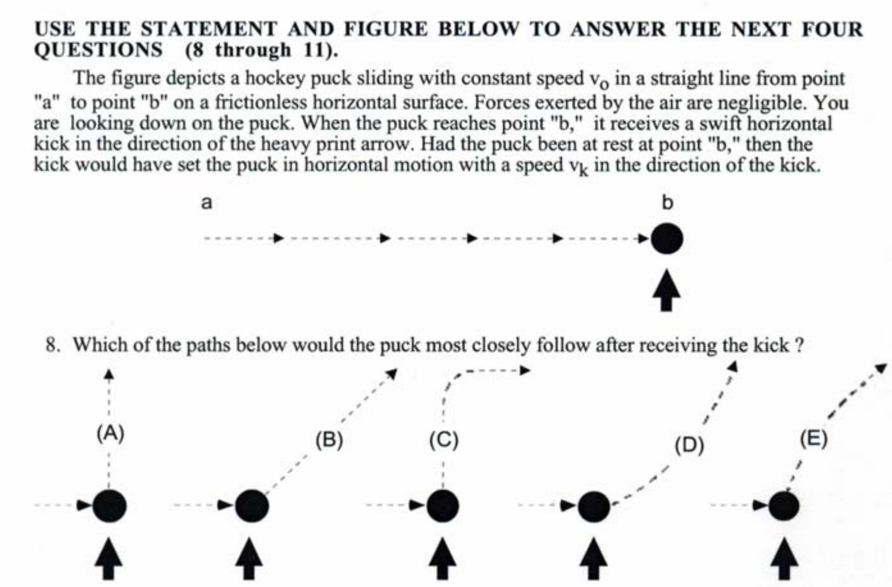


Figure 2. An FCI Problem that tests students' knowledge of gravity, but involves no horizontal motion(from Hestenes et al., 1992).

situation of the kind described in the task model and describable by the task model variables. For multiple-choice tasks like those on the FCI, options are constructed by giving one or more predictions or explanations each that are consistent with Newtonian reasoning, with Galilean impetus reasoning, and with nonscientific (e.g., Aristotelian) reasoning. As illustrated in Section 6, one can reason by analogy to another science domain in which students have misconceptions to create one or more new templates from which to develop an entirely new measure.

H Problems. A typical H Problem from the FCI is shown in Figure 3. The template entitled Force Problems—Horizontal, refers to the identity, direction, amount, and duration of force applied; the mass and identity of the object being moved; whether an illustration is included; the speed and direction of the object being moved; and the time period of interest (before, during, or after the force is applied). (See Appendix B for a screen shot of the Template and Task Model Variables from the Force Problems—Horizontal template.) Note that seven task model variables are common between the G Problem and H Problem templates: the


USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT FOUR QUESTIONS (8 through 11).

The figure depicts a hockey puck sliding with constant speed v_o in a straight line from point "a" to point "b" on a frictionless horizontal surface. Forces exerted by the air are negligible. You are looking down on the puck. When the puck reaches point "b," it receives a swift horizontal kick in the direction of the heavy print arrow. Had the puck been at rest at point "b," then the kick would have set the puck in horizontal motion with a speed v_k in the direction of the kick.

8. Which of the paths below would the puck most closely follow after receiving the kick ?

Figure 3. An FCI Problem that tests students' knowledge of horizontal motion, but does not test knowledge of gravity (from Hestenes et al., 1992).

direction of motion, duration, friction, the identity of the force applied, whether an illustration is included, the mass of the object being moved, and the time period of interest. For example, the problem shown in Figure 3 involves an instantaneous "kick" of unspecified force delivered at right angles to the direction of motion of a hockey puck; the puck's speed and mass are not specified (although we can assume it is lighter than e.g., a rocket); we are told to imagine a frictionless situation; and the problem is illustrated.

Task model variables that are particularly important in eliciting misconceptions are the direction and duration of the force. Table 1 shows examples of how certain combinations of task model variables for a Newton's Third Law problem ("for every action there is an equal and opposite reaction") can tend to elicit specific conceptions or misconceptions (see Hammer & Elby, 2003).

H Problems in the FCI ask about a hockey puck that receives a kick, a rocket in space whose engine turns on, a car that is pushing a truck, a truck that collides with a car, a woman pushing a box, and two students on office chairs who push away from each other. This template likewise would help a researcher create more FCI-like H Problems or to create a new measure.

Table 1

Conceptions That Tend to be Provoked by Combinations of Task Model Variables

Object that gets hit	Object that does the hitting	
	Heavy	Fast
Light	More force from the heavier object	———
Heavy	Correct Newtonian conception	———
Slow	———	More force from the faster object
Fast	———	Correct Newtonian conception

Force Problems—Horizontal and Gravity. The third task model that we created was for more complex problems that require students to apply their knowledge of both horizontal motion and the effect of gravity. A typical H & G Problem from the FCI is shown in Figure 4. The template entitled *Force Problems—Horizontal and Gravity* refers to the identity, direction, and amount of force applied; the mass and

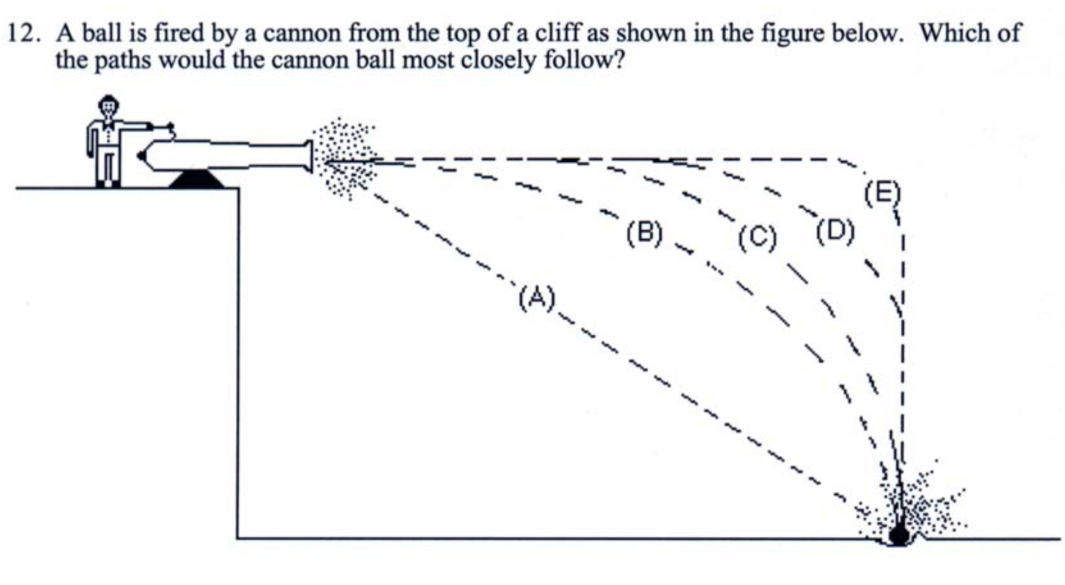


Figure 4. An FCI Problem that tests students' knowledge of both horizontal motion and gravity (from Hestenes et al., 1992).

identity of the object being moved; whether an illustration is included; the direction of motion and the shape of the path of motion of the object being moved. (See Appendix C for a screen shot of the Template and Task Model Variables from the *Force Problems—Horizontal and Gravity* template.) Note that several task model variables from the G Problem and H Problem templates are missing from this more complex template: the duration of force applied, the speed and direction of object; the mass, size, and identity of the object; and the time period of interest (before, during, or after force is applied). For example, the problem shown in Figure 4 involves the instantaneous force of a cannon (the amount of force is not specified) shooting a (presumably heavy) cannonball; the shot is made in a forward direction parallel to the ground; the question asks examinees to determine the shape of the path of motion; and an illustration is included.

G & H problems in the FCI ask about a cannonball shot out of a cannon, balls rolling off a cliff, a bowling ball dropped from a flying airplane, and a tennis ball hit into the wind. Similar to the other two templates, this template would allow a researcher to create more FCI-like G & H Problems or to create a new measure.

6.0 Illustrations Concerning Gas Laws

A second scientific domain in which students have been shown to hold misconceptions is the gas laws—systematic relationships among the pressure, volume, and temperature of gases in closed containers. Some misconceptions that have been observed include the notion that in a partial vacuum gases "rise to the top" or "sink to the bottom" of their container (see Benson, Wittrock, & Baur, 1993; Lin, Cheng, & Lawrenz, 2000; Mas, Perez, & Harris, 1987; Meheut, 1997). On the basis of several interview studies, Novick and Nussbaum (1978, 1981) identified five core beliefs that make up a scientific conception of gas behavior—for example, that gases are made up of particles, that there is empty space between the particles, that the particles are uniformly distributed in a closed container, and that the particles are in constant motion.

Novick and Nussbaum (1981) constructed the Test About Particles in a Gas (TAP), a noncomputational measure of students' conceptual knowledge about gas behavior. It is an 8-question measure that combines multiple-choice questions with drawing tasks and other constructed response questions that are scored by coders into conceptual categories. A sample problem from the TAP is shown in Figure 5.

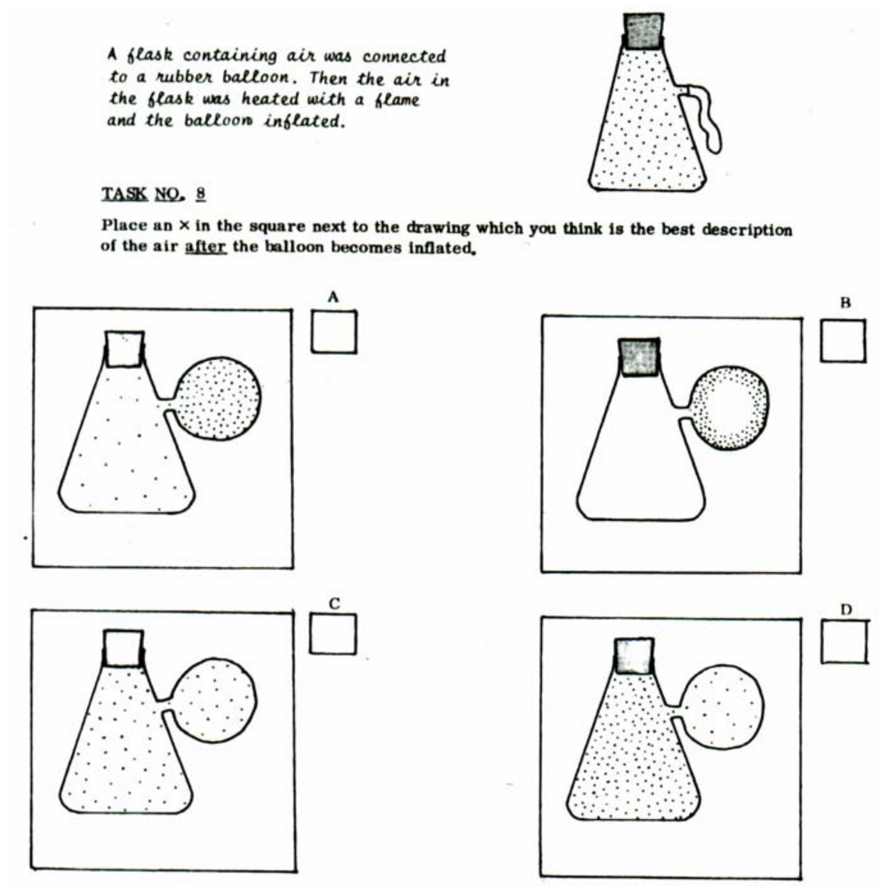


Figure 5. A TAP Problem that tests students' knowledge of the molecular theory of gases (from Novick & Nussbaum, 1981).

Like the FCI, the TAP is designed to reveal students' conceptions about how gases behave, not to test their ability to compute typical Gas Law Problems. The typical Gas Law Problems below all involve computation:

1. Determine the number of grams of carbon dioxide in a 450.6 mL tank at 1.80 atm and minus 50.5 °C. Determine the number of grams of oxygen that the same container will contain under the same temperature and pressure.
2. 1.09 g of H_2 is contained in a 2.00 L container at 20.0 °C. What is the pressure in this container in mm Hg?
3. If 9.006 grams of a gas are enclosed in a 50.00 liter vessel at 273.15 K and 2.000 atmospheres of pressure, what is the molar mass of the gas? What gas is this?

From Diamond Bar High School, Walnut Valley Unified School District, CA
http://dbhs.wvusd.k12.ca.us/GasLaw/WS-Ideal.html

Student model. As with the FCI, we propose that students can be seen as being in multiple belief states, each with a certain propensity. For example, students might simultaneously believe that particles in a gas are uniformly distributed in a closed container and also believe that particles are non-uniformly distributed, each with a certain propensity. As with the FCI, the original student model was a simple univariate model (students are either scientific or nonscientific).

Evidence model. Because the TAP includes both multiple-choice and constructed response questions, its evidence model is more complicated than those of the three FCI templates.

Evaluation procedures and evidence rules. The evidence rules for multiple-choice items on the TAP are analogous to those for the FCI evidence models—evidence about the student corresponds to which multiple-choice option was selected. For constructed response items, however, the response must be coded by a coder into *a priori* categories that correspond to model classifications that are mapped onto Novick and Nussbaum's (1981) five principles—that is, correct understandings or misunderstandings that are keyed to one or more of the five targeted principles. For example, when students make a drawing of gas in a container, coders would have to rate the drawing as showing a particulate vs. continuous conception of gases (see also Benson et al., 1993).

Measurement model. We propose that the Andersen-Rasch measurement model (Andersen, 1973) could likewise be applied to the TAP. Each student would be modeled as being in several belief states, each with a certain propensity; each problem would also be seen as provoking certain belief states, each with a certain propensity. As with the FCI, the original statistical model for the TAP was a simple additive model.

Task model. The task model for TAP questions has two variables in common with the FCI: whether an illustration is used and the nature of the substances involved in the problem. Many of the task model variables are similar to those that would be used in a conventional calculation problem, for example, pressure, volume, and temperature in a closed system; however, TAP problems never give numerical quantities (similar to FCI problems). As in the FCI, students make predictions about or explanations of gas behavior. Interestingly, Novick and Nussbaum (1981) found that problem number 8 (shown in Figure 5) was often answered incorrectly, even by students who demonstrated a correct conception of uniform distribution of gases in a rigid container. An additional task model variable is therefore the rigidity of the container (e.g., a rigid glass container vs. a flexible balloon). TAP problems also vary the response format—for example, multiple choice vs. drawing. As with the FCI, each TAP problem only taps one of the five scientific gas principles.

7.0 A General Design Pattern

We see the three FCI templates (H, G, and H & G), as well as the TAP, as being instantiations of a general type of scientific reasoning that Stewart and colleagues refer to as "Model Using" (Stewart & Hafner, 1994; Stewart, Hafner, Johnson, & Finkel, 1992; See Appendix E). In model using, students reason through a given model about a scientific problem. These problems do not present anomalous data that challenge students' existing models; rather, they provide practice on applying targeted scientific models. In model using, students may consolidate their understanding of a model and how to apply it to problems.

Model using is itself an instantiation of a general class of scientific reasoning involving models, called model-based reasoning. In addition to model using, Stewart has identified Model Elaboration, and Model Revising (Stewart & Hafner, 1994). Model elaboration and model revising differ from model using primarily in that these types of reasoning involve anomalous data that do not fit students' existing models. Model using, on the other hand, involves practice in applying a model to a situation that does not contradict the targeted model.

We see model using as a flexible Design Pattern that can encompass student reasoning with both correct (i.e., scientific) and partially correct (e.g., diSessa's [1993] "p-prims") models. Model using requires knowledge of a model, but especially conditional knowledge about when the model may be applied to a situation (see also Larkin & Simon, 1987). In addition, students need domain-specific or general knowledge in order to use models. Students can use knowledge of the model and their domain-specific knowledge together to make predictions about and explanations of phenomena that the model applies to. For example, given a model of gravity as a force that pulls objects downward and general knowledge about elevators (e.g., that they are heavy, that they are attached to cables that are moved by pulleys), students can reason through their gravity model to make predictions about or explanations of the motion of an elevator (e.g., they could predict that if the elevator cable were cut, the elevator would fall due to gravity).

The FCI and TAP templates, in addition to using Model Using, also depend on a pre-existing body of research that has identified student misconceptions for a particular domain. These are required in order to construct distracters that will appeal to students who do not hold scientific conceptions. For example, the TAP distracters were chosen based on Novick and Nussbaum's (1978) interview study; the FCI distracters were built from a large body of prior studies by researchers such as Clement (1989), Frederiksen and White (1988), Hunt and Minstrell (1994), McCloskey (1983), and others.

8.0 Final Comments

The ultimate goal of the design pattern and task templates is that researchers, teachers, and test developers can have worked-out samples of assessments that are theory-based, psychometrically sound, less burdensome than a "reinvent the wheel" approach, and make optimal use of technology. For example, the Model Using design pattern and the four associated templates developed here are based in research into students' misconceptions in science—one application of the novice-expert paradigm in cognitive psychology. They are able to take a significant part of the assessment development burden off of task designers, whether they choose to use the templates as is, or adapt them.

Tasks developed from these templates are grounded in a branch of cognitive research in science learning. They could be developed either for informal use in the classroom or for larger scale and more formal use with psychometric models. In assessment systems that follow this latter route, the templates' theory-based student models and aligned evidence models are psychometrically sound. They take advantage of sophisticated statistical models (e.g., Rasch models, Bayes nets) that need not be developed afresh, and the details of which can be made invisible to the task designer and end user. And finally, since they can be expressed in transportable form, they can be provided in ways that are simple to share and access via the Web, and to use in systems that take advantage of automated student recordkeeping, adaptivity, and other features of technology.

References

Adams, R., Wilson, M. R., & Wu, M. (1997). Multilevel item response models: An approach to errors in variables regression. *Journal of Educational and Behavioral Statistics, 22,* 47-76.

Andersen, E. B. (1973). *Conditional inference and models for measuring.* Copenhagen: Danish Institute for Mental Health.

Andersen, E. B. (1995). Polytomous Rasch models and their estimation. In G. H. Fischer, & I. W. Molenaar (Eds.), *Rasch models: Foundations, recent developments, and applications* (pp. 271-292). New York: Springer-Verlag.

Bao, L., & Redish, E. F. (2001). *Model analysis: Assessing the dynamics of student learning.* Retrieved November 25, 2003, from http://www.physics.ohio-state.edu/~lbao/archive/papers/MAPre_01-0805.pdf

Benson, D. L., Wittrock, M. C., & Baur, M. E. (1993). Students' preconceptions of the nature of gases. *Journal of Research in Science Teaching, 30,* 587-597.

Bishop, B. A., & Anderson, C. W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching, 27,* 415-427.

Brown, J. S., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science, 4,* 379-426.

Breuleux, A. (1991). The analysis of writers' think-aloud protocols: Developing a principled coding scheme for ill-structured tasks. In G. Denhiere & J. P. Rossi (Eds.), *Text and text processing* (pp. 333-362). Amsterdam: Elsevier.

Charness, N., & Schultetus, R. (1999). Knowledge and expertise. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay, & M. T. H. Chi (Eds.), *Handbook of applied cognition* (pp. 57-81). New York: John Wiley.

Chase, W. G. (1982). Spatial representations of taxi drivers. In D. R. Rogers & J. A. Sloboda (Eds.), *Acquisition of symbolic skills* (pp. 391-405). New York: Plenum Press.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.

Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121-152.

Clement, J. (1983). A conceptual model discussed by Galileo and used intuitively by physics students. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299-324). Hillsdale, NJ: Erlbaum.

Clement, J. (1989). Learning via model construction and criticism. In G. Glover, R. Ronning, & C. Reynolds (Eds.), *Handbook of creativity: Assessment, theory, and research* (pp. 341-381). New York: Plenum Press.

Cornely-Moss, K. A. (1995). Kinetic theory of gases. *Journal of Chemical Education, 72,* 715-716.

Dalton, B., Morocco, C. C., Tivnan, T., & Mead, P. L. R. (1997). Supported inquiry science: Teaching for conceptual change in urban and suburban science classrooms. *Journal of Learning Disabilities, 30,* 670-684.

de Groot, A. (1946/1978). *Thought and choice in chess.* The Hague: Mouton.

diSessa, A. A. (1993). Toward and epistemology of physics. *Cognition and Instruction, 10,* 105-225.

Ericsson K. A. (Ed). (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games.* Hillsdale, NJ: Erlbaum.

Ericsson, K. A., & Charness, N. (1997). Cognitive and developmental factors in expert performance. In P. J. Feltovich, K. M. Ford, & R. R. Hoffman (Eds.), *Expertise in context: Human and machine* (pp. 3-41). Menlo Park, CA: AAAI Press.

Ericsson, K. A., & Polson, P. G. (1988). A cognitive analysis of exceptional memory for restaurant orders. In M. Chi., R. Glaser, & M. Farr (Eds.), *The nature of expertise* (pp. 23-70). Hillsdale NJ: Erlbaum.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits of the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1-38). New York: Cambridge University Press.

Frederiksen, J. R., & White, B. Y. (1988). Implicit testing within an intelligent tutoring system. *Machine-Mediated Learning, 2,* 351-372.

Gardner, H. (1999). *The disciplined mind.* New York: Simon & Schuster.

Hammer, D., & Elby, A. (2003). Tapping epistemological resources for learning physics. *Journal of the Learning Sciences, 12,* 53-90.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30,* 141-151.

Huang, C.-W. (2003). *Psychometric analyses based on evidence-centered design and cognitive science of learning to explore students' problem-solving in physics.* Unpublished doctoral dissertation, University of Maryland, College Park, Department of Measurement, Statistics, and Evaluation.

Hunt, E., & Minstrell, J. (1994). A cognitive approach to the teaching of physics. In K. McGilly (Ed.), *Classroom lessons: Integration cognitive theory and classroom practice* (pp. 51-74). Cambridge, MA: MIT Press.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, 208,* 1335-1342.

Larkin, J., & Simon, H. (1987) Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Lin, H., Cheng, H., & Lawrenz, F. (2000). The assessment of students and teachers' understanding of gas laws. *Journal of Chemical Education, 77,* 715-716.

Lundeberg, M. A. (1987). Metacognitive aspects of reading comprehension: Studying understanding in legal case analysis. *Reading Research Quarterly, 22,* 407-432.

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics, 20,* 193-232.

Mas, C. J. F., Perez, J. H., & Harris, H. H. (1987). Parallels between adolescents' conception of gases and the history of chemistry. *Journal of Chemical Education, 64,* 616-618.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299-324). Hillsdale, NJ: Erlbaum.

Meheut, M. (1997). Designing a learning sequence about a prequantitative kinetic model of gases: The parts played by questions and by a computer-simulation. *International Journal of Science Education, 19,* 647-660.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Education Researcher, 32*(2), 13-23.

Michael, J. A. (1998). Students' misconceptions about perceived physiological responses. *Advances in Physiology education, 19,* S90-S98.

Mislevy, R. J., Chudowsky, N., Draney, K., Fried, R., Gaffney, T., Haertel, G., et al. (2003). *Design patterns for assessing science inquiry. Principled Assessment Designs for Inquiry* (PADI Tech. Rep. 1). Menlo Park, CA: SRI International. Available 8 July 2003 from http://padi.sri.com/downloads/ PADI_DesignPatterns.pdf

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*, 3-67.

Niaz, M. (2000). Gases as idealized lattices: A rational reconstruction of students' understanding of the behavior of gases. *Science and Education, 9*, 279-287.

Novick, S., & Nussbaum, J. (1978). Junior high school pupils' understanding of the particulate nature of matter: An interview study. *Science Education, 62*, 273-281.

Novick, S., & Nussbaum, J. (1981). Pupils' understanding of the particulate nature of matter: A cross-age study. *Science Education, 65,* 187-196.

Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 93-125). New York: Cambridge University Press.

Peskin, J. (1998). Constructing meaning when reading poetry: An expert-novice study. *Cognition and Instruction, 16,* 235-263.

Riconscente, M., Mislevy, R. J., & Hamel. L. (2004). *An introduction to PADI task templates* (PADI Tech. Rep. 2). Menlo Park, CA: SRI International.

Rouet, J.-F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology, 88,* 478-493.

Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multi-dimensionality, and immediacy. *American Educational Research Journal, 28*, 63-88.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8,* 481-520

Steinberg, L. S., & Gitomer, D. G. (1996). Intelligent tutoring and assessment built on an understanding of a technical problem-solving task. *Instructional Science, 24,* 223-258.

Stewart, J., & Hafner, R. (1994). Research on problem solving: Genetics. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 284-300). New York: Macmillan.

Stewart, J., Hafner, R., Johnson, S., & Finkel, E. (1992). Science as model building: Computers and high-school genetics. *Educational Psychologist, 27,* 317-36.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, *24*, 535-585.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, *16*, 3-18.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83*, 73-87.

Appendix A

PADI Template for Force Problems—Gravity Only

The PADI template for “Force Problems—Gravity Only” is reproduced below. Templates include information about the student model, the measurement model, and the task model. The **student model** information is contained in the “Student Model Summary” section and the link to the student model. The summary explains the multivariate student model proposed above, while the linked Student Model defines model identification and the covariance matrix.

Measurement model information is contained in the “Evaluation Procedures Summary” (evidence rules) and “Measurement Model Summary” (statistical model). The evaluation procedures explain the mapping of the multiple-choice items to Newtonian, Galilean, and non-scientific conception categories. The measurement model summary explains the Andersen-Rasch model discussed above.

Task model information is contained in the sections entitled “Activities,” “Activity Sequencing,” “Template-level Task Model Variables,” and “Task Model Variable Settings.” The primary activity for this template is “Make explanations and predictions from a physical situation,” which is linked to the template. The task model variables described above are linked to this particular template, with template-specific comments (e.g., the “Direction of Motion” TMV in this template is always vertical).

In addition, the template includes links to the model using Design Pattern, relevant research, Web resources (e.g., the Web site for the FCI), and other print resources, such as the Bao and Redish (2001) paper.

PADI Design System Home

Logged in as: rmislevy Logout

View Template 318. "Force Problems--Gravity Only"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Force Problems--Gravity Only	
Summary Edit	This template is for multiple-choice conceptual Newtonian force problems that involve gravity, but not both gravity and horizontal motion; also, they do not include circular motion.	This template can produce a measure like a subset of Newtonian force problems from the Force Concept Inventory (FCI).
Type Edit	Abstract Template	
Student Model Summary Edit	Student is seen as being in several belief states simultanously (e.g., Newtonian, Galilean, and 'folk'), each with a certain propensity.	In some ways analogous to the Siegler (1976) balance beam study.
Student Models Edit	FCI-ish student model Propensities to respond to tasks in a certain class (e.g., FCI and FCME items) in terms of conceptio...	
Measurement Model Summary Edit	Measurement model is multivariate, a la Andersen/Rasch model. Original measurement model was univariate.	Measurement model is an Andersen/Rasch model (Andersen, 1995) being tested by Kevin Huang at UMCP. It can be written as a special case of the MRCMLM (Adams, Wilson, & Wu, 1997). Previous models include Bao and Redish (2001) or simple summary scores.
Evaluation Procedures Summary Edit	Answer key--one correct answer per question.	Correct answer always corresponds to the result of reasoning from a Newtonian conception. (In principle, could have open-ended answers, which would be mapped into model categories--e.g., Minstrell, 1989)

Continued

Work Product Summary Edit	Answers to multiple-choice problems are main work product.	Multiple choice responses have been constructed so that each one a priori maps to exactly one of the model classifications. Correct answers apply Newtonian principles correctly to the problem. Incorrect answers reflect non-Newtonian principles; specifically, misconceptions identified in prior literature, which might be Galilean, Aristotelian, or non-scientific. The list of Newtonian concepts tested in the FCI is listed in Hestenes, Wells & Swackhamer (1992), Table 1. The list of common misconceptions used to construct distractors is listed in Table 2 of the same paper.
Task Model Variable Summary Edit	Task model variables include direction of motion (up, down, or no motion); time elapsed; mass of object; forces other than gravity.	Downward force always = g = 32 ft/sec^2, friction always = 0, and Motion is always in a vertical plane in this subset of problems. Ex: FCI Q1 direction of motion = down time elapsed = (>, =, <) mass of object = same for both objects Ex: FCI Q17 direction of motion = up forces other than gravity = force of cable opposite to force of gravity. Irrelevant to this Q: time elapsed, mass of object
Task Model Variable Settings Edit	View	
Presentation Environment Requirements Edit	Paper-and-pencil measure.	Could easily be adapted for computer administration.
Materials and Presentation Settings Edit	View	
Activities Summary Edit		
Activities Edit	Make explanations and predictions from a physical situation Given a physical situation with some underlying regularities (e.g., a Newtonian law or gas law), stu...	

Continued

Activity Sequencing	Edit	Sets of questions relating to a single situation may be presented or answered in any order. Ordinarily all questions on the FCI are administered at one sitting (it is commonly used as a pretest-posttest measure in university introductory physics courses).
Template-level Task Model Variables	Edit	Direction of motion The path taken by an object in motion e.g., Answer choices refer to the path of motion of an ob... Duration e.g., in FCI-type problem, is a force applied instantaneously or over a period of time. Friction Amount of friction between two objects such as a puck and ice, a rope and a pulley, a ball and air. Identity of horizontal force The identity/ies of a force(s) which act(s) on an object(s) for an instant or over a period of time,... Illustrations Whether or not an illustration (e.g., line drawing, photograph, video clip) is part of the task. Mass e.g., Mass of object to which force is applied. Objects Objects on which forces act. They may vary in size, mass, animate/inanimate, etc. Speed e.g. The speed of an object at the moment when a force is applied to it. Time period of interest Past, present, or future. e.g., Question asks about motion or speed while a force is applied vs. af...
Tools for Examinee	Edit	Pen or pencil.
Exemplars	Edit	Force Concept Inventory (FCI) A multiple-choice measure of students' conceptual knowledge of forces, including Newton's 1st-3rd la...
Educational Standards	Edit	
Design Patterns	Edit	Model Using This design pattern generates tasks where the student has to reason through a given model using data...
I am a part of	Edit	SRI Group TOMS 1/27
These are parts of me	Edit	
Online resources	Edit	FCI available at http://modeling.la.asu.edu/R&E/Research.html
References	Edit	Bao, L., & Redish, E.F. (2001). Model analysis: Assessing the dynamics of student learning. — Available on the World Wide Web at http://www.physics.ohio-state.edu/~lbao/archive/papers/MAPre_01-0805.pdf Wang (2003)

PADI Design System Home

Logged in as: rmislevy Logout

View Task Model Variable 314. "Direction of motion"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Direction of motion	
Summary Edit	The path taken by an object in motion e.g., Answer choices refer to the path of motion of an object (perpendicular vs. acute angle)	e.g., for FCI-type questions
TMV Type Edit	Free-form text entry	
TMV Category (possible value) Edit	Continuing in a straight line	these are samples of possible values for this variable
	Down	
	No motion	
	Perpendicular turn	
	Turns at an angle	
	Up	
Online resources Edit		
References Edit		

View in XML

Design System Home

Logged in as: rmislevy   Logout

View Task Model Variable 316. "Duration"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Duration	
Summary Edit	e.g., in FCI-type problem, is a force applied instantaneously or over a period of time.	e.g., for FCI-type problem. Must be greater than zero. Units of measurement and measurement instrument may be issues
TMV Type Edit	Free-form text entry	
TMV Category (possible value) Edit	For 5 seconds	these are samples of possible values for this variable
	For a time	
	Instantaneously	
Online resources Edit		
References Edit		

View in XML

PADI [Design System Home](#)

Logged in as: rmislevy Logout

View Activity 288. "Make explanations and predictions from a physical situation"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Make explanations and predictions from a physical situation	
Summary Edit	Given a physical situation with some underlying regularities (e.g., a Newtonian law or gas law), students explain what is happening in the situation or make a prediction about what will happen.	Activities that are tested by assessments such as the Force Concept Inventory (FCI) and Test About Particles in a Gas (TAP).
Measurement Models Edit	Andersen/Rasch measurement model fragment Multinomial generalization of Rasch model.	
Evaluation Procedures Edit	Map multiple-choice answers onto conceptual model e.g., FCI answers mapped onto Newtonian, Aristotelian, and non-scientific models.	
Work Products Edit	Choice from profferred choices The student is to make a selection from among presented choices Right/wrong multiple choice items a...	
Materials and Presentation Edit	Problem 1-1 text Two objects start at the same position, end at different positions, and travel in the same direction...	
	Problem 1-2 text Two objects start at different positions, end at same position, and travel in the same direction.	
	Problem answer choices For problems that present answer choices in either text or illustrations.	
	Problem illustration For problems that include an illustration to be used in solving the problem.	
	Problem text For problems that include text (e.g., a question, scenario, case study, etc.).	
Presentation Logic Edit		
Task Model Variables Edit		
Design Patterns Edit	Model elaboration One type of problem solving that involves learning is the explanation and use of a given model.	
	Model Using This design pattern generates tasks where the student has to reason through a given model using data...	
Online resources Edit		
References Edit	Refer to Design Pattern (e.g., Model Using)	

View in XML

Design System Home

Logged in as: rmislevy Logout

View Student Model 334. "FCI-ish student model"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	FCI-ish student model	
Summary Edit	Propensities to respond to tasks in a certain class (e.g., FCI and FCME items) in terms of conception models. The conceptions are Newtonian, Galilean, and nonscientific. The model is identified by implicitly fixing all examinees at zero, so there are only two SM variables, namely for the Galilean and nonscientific propensities.	
Distribution Type Edit	Multivariate normal	
Distribution Summary Edit	Bivariate normal for the two SM variables that are explicit	
Covariance Matrix Edit	View	1 0 0 1
Means Matrix Edit	View	
Student Model Variables Edit	Galilean propensity nonscientific propensity	

Design System Home

Logged in as: rmislevy Logout

View Evaluation Procedure (rubric) 290. "Map multiple-choice answers onto conceptual model"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Map multiple-choice answers onto conceptual model	
Summary Edit	e.g., FCI answers mapped onto Newtonian, Aristotelian, and non-scientific models.	
Phase sequencing Edit	One phase	
Evaluation Phases Edit	Model using multiple choice Multiple-choice evaluation option for task templates developed under model-using design pattern such...	

Design System Home

Logged in as: rmislevy Logout

View Evaluation Phase 327. "Model using multiple choice"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Model using multiple choice	
Summary Edit	Multiple-choice evaluation option for task templates developed under model-using design pattern such as FCI-like or TAP-like assessments.	Note that an evaluation phase entitled "model using open-ended" has also been created, but is not yet linked to an activity or template.
Work Products Edit	Choice from profferred choices The student is to make a selection from among presented choices Right/wrong multiple choice items a...	
	Choice of a Multiple Choice Answer The task requires students to read a question and choose the best answer.	
Task Model Variables Edit		
Preceding Evaluation Phase Edit		
Observable variables Edit	Conception category Which conceptual category the examinee's response is in (e.g., Newtonian vs. Galilean for FCI)	
Evaluation Action Edit	Identify which of the options the examinee has chosen, and take its mapping into the conception/ misconception category from the designer-provided coding key	
Evaluation Action Data Edit	Use a list of known scientific conceptions and misconceptions from prior research (e.g., Tables 1 & 2 from Hestenes et al., 1992) to develop the coding key	

Design System Home

Logged in as: rmislevy Logout

View Measurement Model 289. "Andersen/Rasch measurement model fragment"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Andersen/Rasch measurement model fragment	
Summary Edit	Multinomial generalization of Rasch model.	See Andersen (1995)
Type of Measurement Model Edit	(not specified)	measurement model fragment for the Andersen/Rasch model, which posits a 3 categories: Newtonian, Galilean, and "nonscientific". In principle each student and each item would have three parameters also, for their propensities for each category. Under the MRCML, however, identification is achieved by fixing every student's first parameter at zero and every item's first parameter at zero. So we have SM and every item with two parameters--in this case, for the Galilean & nonscientific categories.
Observable Variable Edit	Conception category Which conceputal category the examinee's response is in (e.g., Newtonian vs. Galilean for FCI)	
Student Model Variables Edit	Galilean propensity nonscientific propensity	

Appendix B

PADI Template for Force Problems—Horizontal Only

The PADI template for "Force Problems—Horizontal" has links similar to the links for the template for "Force Problems—Gravity," but has its own Template-level Task Model Variables. Only the top layer of this template is shown.

Design System Home

Logged in as: rmislevy Logout

View Template 276. "Force Problems--Horizontal"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Force Problems--Horizontal	
Summary Edit	This template is for multiple-choice conceptual Newtonian force problems that do not involve gravity or circular motion.	This template can produce a measure like a subset of Newtonian force problems from the Force Concept Inventory (FCI).
Type Edit	Abstract Template	
Student Model Summary Edit	Student is seen as being in several belief states simultanously (e.g., Newtonian, Galilean, and 'folk'), each with a certain propensity.	In some ways analogous to the Siegler (1976) balance beam study.
Student Models Edit	FCI-ish student model Propensities to respond to tasks in a certain class (e.g., FCI and FCME items) in terms of conceptio...	
Measurement Model Summary Edit	Measurement model is multivariate, a la Andersen/Rasch model. Original measurement model was univariate.	Measurement model is an Andersen/Rasch model (Andersen, 1995) being tested by Kevin Huang at UMCP. It can be written as a special case of the MRCMLM (Adams, Wilson, & Wu, 1997). Previous models include Bao and Redish (2001) or simple summary scores.
Evaluation Procedures Summary Edit	Answer key--one correct answer per question.	Correct answer always corresponds to the result of reasoning from a Newtonian conception. (In principle, could have open-ended answers, which would be mapped into model categories--e.g., Minstrell, 1989)

Continued

Work Product Summary Edit	Answers to multiple-choice problems are main work product.	Multiple choice responses have been constructed so that each one a priori maps to exactly one of the model classifications. Correct answers apply Newtonian principles correctly to the problem. Incorrect answers reflect non-Newtonian principles; specifically, misconceptions identified in prior literature, which might be Galilean, Aristotelian, or non-scientific. The list of Newtonian concepts tested in the FCI is listed in Hestenes, Wells & Swackhamer (1992), Table 1. The list of common misconceptions used to construct distractors is listed in Table 2 of the same paper.
Task Model Variable Summary Edit	Task model variables include identity, direction, amount, and duration of force applied; mass and identity of object; speed and direction of object; time period of interest (before, during, or after force is applied).	e.g., FCI Q25 amount of force = "constant horizontal force" mass = "large box" speed after force is applied = "Vo" Friction always = 0 and Motion is always in a horizontal plane in this subset of problems, therefore gravity is never a force. Other variables are implicitly or explicitly set constant (e.g., direction not specified, duration not specified, speed of object before force = 0, time period of interest = while the force is applied, path of motion must be a straight line).
Task Model Variable Settings Edit	View	
Presentation Environment Requirements Edit	Paper-and-pencil measure.	Could easily be adapted for computer administration.
Materials and Presentation Settings Edit	View	
Activities Summary Edit		

Continued

Activities	Make explanations and predictions from a physical situation Given a physical situation with some underlying regularities (e.g., a Newtonian law or gas law), stu...
Activity Sequencing	Sets of questions relating to a single situation may be presented or answered in any order. Ordinarily all questions on the FCI are administered at one sitting (it is commonly used as a pretest-posttest measure in university introductory physics courses).
Template-level Task Model Variables	Amount of force applied Amount of force which one object exerts on another (moving or stationary) object. Direction of force applied The direction in which a force is applied to an object Direction of motion The path taken by an object in motion e.g., Answer choices refer to the path of motion of an ob... Duration e.g., in FCI-type problem, is a force applied instantaneously or over a period of time. Friction Amount of friction between two objects such as a puck and ice, a rope and a pulley, a ball and air. Identity of horizontal force The identity/ies of a force(s) which act(s) on an object(s) for an instant or over a period of time,... Illustrations Whether or not an illustration (e.g., line drawing, photograph, video clip) is part of the task. Mass e.g., Mass of object to which force is applied. Objects Objects on which forces act. They may vary in size, mass, animate/inanimate, etc. Shape of path of motion The shape of the path taken by an object in motion e.g., Answer choices refer to the path of mo... Speed e.g. The speed of an object at the moment when a force is applied to it. Time period of interest Past, present, or future. e.g., Question asks about motion or speed while a force is applied vs. af...
Tools for Examinee	Pen or pencil.

Continued

Exemplars	Edit	Force Concept Inventory (FCI) A multiple-choice measure of students' conceptual knowledge of forces, including Newton's 1st-3rd la...
Educational Standards	Edit	
Design Patterns	Edit	Model Using This design pattern generates tasks where the student has to reason through a given model using data...
I am parent of	Edit	
I am child of	Edit	
I am a part of	Edit	
These are parts of me	Edit	
Online resources	Edit	FCI available at http://modeling.la.asu.edu/R&E/Research.html
References	Edit	Bao, L., & Redish, E.F. (2001). Model analysis: Assessing the dynamics of student learning. — Available on the World Wide Web at http://www.physics.ohio-state.edu/~lbao/archive/papers/MAPre_01-0805.pdf Wang (2003)

View in XML

Appendix C

PADI Template for Force Problems—Horizontal and Gravity

The PADI template for "Force Problems—Horizontal and Gravity" has links similar to links in the FCI templates above, but its own Template-level Task Model Variables. Only the top layer of this template is shown.

PADI Design System Home

View Template 322. "Force Problems--Horizontal and Gravity"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Force Problems--Horizontal and Gravity	
Summary Edit	This template is for multiple-choice conceptual Newtonian force problems that simultaneously involve horizontal motion and the downward pull of gravity, but not circular motion.	This template can produce a measure like a subset of Newtonian force problems from the Force Concept Inventory (FCI). Force Problems--Horizontal and Force Problems--Gravity are special cases of this template.
Type Edit	Abstract Template	
Student Model Summary Edit	Student is seen as being in several belief states simultanously (e.g., Newtonian, Galilean, and 'folk'), each with a certain propensity.	In some ways analogous to the Siegler (1976) balance beam study.
Student Models Edit		
Measurement Model Summary Edit	Measurement model is multivariate, a la Andersen/Rasch model. Original measurement model was univariate.	Measurement model is an Andersen/Rasch model (Andersen, 1995) being tested by Kevin Huang at UMCP. It can be written as a special case of the MRCMLM (Adams, Wilson, & Wu, 1997). Previous models include Bao and Redish (2001) or simple summary scores.
Evaluation Procedures Summary Edit	Answer key--one correct answer per question.	Correct answer always corresponds to the result of reasoning from a Newtonian conception. (In principle, could have open-ended answers, which would be mapped into model categories--e.g., Minstrell, 1989)

Continued

Work Product Summary Edit	Answers to multiple-choice problems are main work product.	Multiple choice responses have been constructed so that each one a priori maps to exactly one of the model classifications. Correct answers apply Newtonian principles correctly to the problem. Incorrect answers reflect non-Newtonian principles; specifically, misconceptions identified in prior literature, which might be Galilean, Aristotelian, or non-scientific. The list of Newtonian concepts tested in the FCI is listed in Hestenes, Wells & Swackhamer (1992), Table 1. The list of common misconceptions used to construct distractors is listed in Table 2 of the same paper.
Task Model Variable Summary Edit	Task model variables include identity, direction, and amount of force applied; released from moving or still object; mass, size, and identity of object; horizontal distance traveled; path of motion.	e.g., FCI Q12 identity of force = cannon blast and gravity simultaneously direction of force = horizontal and downward simultaneously amount of force = cannon blast and g = 32 ft/sec^2 released from moving or still object = still mass = implied heavy size = implied medium sized identity = cannonball path of motion = TBD (parabola) Irrelevant: horizontal distance traveled
Task Model Variable Settings Edit	View	
Presentation Environment Requirements Edit	Paper-and-pencil measure.	Could easily be adapted for computer administration.
Materials and Presentation Settings Edit	View	
Activities Summary Edit		

Continued

Activities	Edit	Make explanations and predictions from a physical situation Given a physical situation with some underlying regularities (e.g., a Newtonian law or gas law), stu...
Activity Sequencing	Edit	Sets of questions relating to a single situation may be presented or answered in any order. Ordinarily all questions on the FCI are administered at one sitting (it is commonly used as a pretest-posttest measure in university introductory physics courses).
Template-level Task Model Variables	Edit	Amount of force applied Amount of force which one object exerts on another (moving or stationary) object. Direction of force applied The direction in which a force is applied to an object Direction of motion The path taken by an object in motion e.g., Answer choices refer to the path of motion of an ob... Friction Amount of friction between two objects such as a puck and ice, a rope and a pulley, a ball and air. Identity of horizontal force The identity/ies of a force(s) which act(s) on an object(s) for an instant or over a period of time,... Illustrations Whether or not an illustration (e.g., line drawing, photograph, video clip) is part of the task. Mass e.g., Mass of object to which force is applied. Objects Objects on which forces act. They may vary in size, mass, animate/inanimate, etc. Shape of path of motion The shape of the path taken by an object in motion e.g., Answer choices refer to the path of mo...
Tools for Examinee	Edit	Pen or pencil.
Exemplars	Edit	Force Concept Inventory (FCI) A multiple-choice measure of students' conceptual knowledge of forces, including Newton's 1st-3rd la...
Educational Standards	Edit	

Appendix D

PADI Template for Gas Law Problems

The PADI template for "Gas Law Problems" reproduced below has its own specific information about the student model, the measurement model (e.g., mapping student-constructed responses to predetermined misconception categories), and task model (e.g., for constructing both multiple-choice and constructed-response items), as well as references to the TAP measure, and other print resources. The same Activity used to structure the FCI Force tasks (displayed in Appendix A) is re-used in this template. Only the top layer of the template is shown.

Design System Home   Logged in as: mmislevy   Logout

View Template 322. "Force Problems--Horizontal and Gravity"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Force Problems--Horizontal and Gravity	
Summary Edit	This template is for multiple-choice conceptual Newtonian force problems that simultaneously involve horizontal motion and the downward pull of gravity, but not circular motion.	This template can produce a measure like a subset of Newtonian force problems from the Force Concept Inventory (FCI). Force Problems--Horizontal and Force Problems--Gravity are special cases of this template.
Type Edit	Abstract Template	
Student Model Summary Edit	Student is seen as being in several belief states simultanously (e.g., Newtonian, Galilean, and 'folk'), each with a certain propensity.	In some ways analogous to the Siegler (1976) balance beam study.
Student Models Edit		
Measurement Model Summary Edit	Measurement model is multivariate, a la Andersen/Rasch model. Original measurement model was univariate.	Measurement model is an Andersen/Rasch model (Andersen, 1995) being tested by Kevin Huang at UMCP. It can be written as a special case of the MRCMLM (Adams, Wilson, & Wu, 1997). Previous models include Bao and Redish (2001) or simple summary scores.
Evaluation Procedures Summary Edit	Answer key--one correct answer per question.	Correct answer always corresponds to the result of reasoning from a Newtonian conception. (In principle, could have open-ended answers, which would be mapped into model categories--e.g., Minstrell, 1989)

Continued

Work Product Summary Edit	Answers to multiple-choice problems are main work product.	Multiple choice responses have been constructed so that each one a priori maps to exactly one of the model classifications. Correct answers apply Newtonian principles correctly to the problem. Incorrect answers reflect non-Newtonian principles; specifically, misconceptions identified in prior literature, which might be Galilean, Aristotelian, or non-scientific. The list of Newtonian concepts tested in the FCI is listed in Hestenes, Wells & Swackhamer (1992), Table 1. The list of common misconceptions used to construct distractors is listed in Table 2 of the same paper.
Task Model Variable Summary Edit	Task model variables include identity, direction, and amount of force applied; released from moving or still object; mass, size, and identity of object; horizontal distance traveled; path of motion.	e.g., FCI Q12 identity of force = cannon blast and gravity simultaneously direction of force = horizontal and downward simultaneously amount of force = cannon blast and g = 32 ft/sec^2 released from moving or still object = still mass = implied heavy size = implied medium sized identity = cannonball path of motion = TBD (parabola) Irrelevant: horizontal distance traveled
Task Model Variable Settings Edit	View	
Presentation Environment Requirements Edit	Paper-and-pencil measure.	Could easily be adapted for computer administration.
Materials and Presentation Settings Edit	View	
Activities Summary Edit		

Continued

Activities	Make explanations and predictions from a physical situation Given a physical situation with some underlying regularities (e.g., a Newtonian law or gas law), stu...
Activity Sequencing	Sets of questions relating to a single situation may be presented or answered in any order. Ordinarily all questions on the FCI are administered at one sitting (it is commonly used as a pretest-posttest measure in university introductory physics courses).
Template-level Task Model Variables	Amount of force applied Amount of force which one object exerts on another (moving or stationary) object. Direction of force applied The direction in which a force is applied to an object Direction of motion The path taken by an object in motion e.g., Answer choices refer to the path of motion of an ob... Friction Amount of friction between two objects such as a puck and ice, a rope and a pulley, a ball and air. Identity of horizontal force The identity/ies of a force(s) which act(s) on an object(s) for an instant or over a period of time,... Illustrations Whether or not an illustration (e.g., line drawing, photograph, video clip) is part of the task. Mass e.g., Mass of object to which force is applied. Objects Objects on which forces act. They may vary in size, mass, animate/inanimate, etc. Shape of path of motion The shape of the path taken by an object in motion e.g., Answer choices refer to the path of mo...
Tools for Examinee	Pen or pencil.
Exemplars	Force Concept Inventory (FCI) A multiple-choice measure of students' conceptual knowledge of forces, including Newton's 1st-3rd la...
Educational Standards	

Continued

Design Patterns [Edit]	Model Using This design pattern generates tasks where the student has to reason through a given model using data...	
I am parent of [Edit]		
I am child of [Edit]		
I am a part of [Edit]		
These are parts of me [Edit]		
Online resources [Edit]	FCI available at http://modeling.la.asu.edu/R&E/Research.html	
References [Edit]	Bao, L., & Redish, E.F. (2001). Model analysis: Assessing the dynamics of student learning.	Available on the World Wide Web at http://www.physics.ohio-state.edu/~lbao/archive/papers/MAPre_01-0805.pdf
	Wang (2003)	

View in XML

Appendix E

PADI Design Pattern for Model Using

The PADI “Model Using” design pattern is reproduced below. It includes a summary, KSAs (knowledge, skills, and abilities), potential observations, potential work products, and characteristic and variable features of the design pattern. The KSAs include both focal KSAs (e.g., the importance of conditional knowledge discussed above) and additional KSAs (e.g., general knowledge). Potential observations and work products are very high level statements that are operationalized in the templates as student model and task model variables (e.g., Work Products in the Activity). Characteristic and variable features are likewise abstract statements that are operationalized in the templates as task model variables (e.g., Task Model Variables).

Design System Home

Logged in as: rmislevy Logout

View Design Pattern 274. "Model Using"

Duplicate

Delete (remove permanently)

Section	Value	Comment
Title Edit	Model Using	
Summary Edit	This design pattern generates tasks where the student has to reason through a given model using data that do not contradict the model.	One type of problem solving that involves learning is the application of a known model to data that do not contradict the model. Students' work is bound by the concept of a known model (or models), so their work includes an understanding of the constraints of the problem.
Rationale Edit	Consolidation of understanding may occur by solving problems in which the data do not conflict with the existing model in use by the solvers. Solvers may consolidate their understanding of the model and how to apply it to problems.	Even though model using does not involve the invention of new objects, processes, or states, it does entail problem solving and is an analogue of much scientific activity.
Focal Knowledge, Skills and Abilities Edit	Being able to reason through the concepts and relationships of a given model.	Models may be correct (conventional scientific) or not (e.g., misconceptions or limited cases [p-prims, facets]), but students do learn how to apply them to problems, and in the process consolidate their understanding of the model.
	Conditional knowledge: knowing when to apply the model.	
Additional Knowledge, Skills and Abilities Edit	Familiarity with task type (e.g., materials, protocols, expectations).	
	Subject-area knowledge (declarative and procedural).	

Continued

Potential observations	Given a model and a situation, making explanations, predictions, or retrodictions reasoned through the model.	This DP is defined by the absence of model modification. In model-using, the model is entirely unproblematic to the student.
	Making and explaining predictions through a model.	
	Mapping out the corresponding elements between a real-world situation and a scientific model.	
Potential work products	Correspondence mapping between elements or relationships of model and real-world situation.	
	Correspondence mapping between elements or relationships of overlapping models.	
	Hypotheses (constructed/selected).	
	Predictions (constructed/selected).	
	Written/Oral Explanation of reasoning.	
Potential rubrics		
Characteristic features	Real-world situation and one or more models appropriate to the situation, for which details of correspondence need to be fleshed out. Addresses correspondence between situation and models, and models with one another. Patterns in the data do not contradict the model.	
Variable features	Is problem context familiar?	
	Must experimental work or supporting research be carried out in order to ground the elaboration?	
	Single model to elaborate, vs. establishing correspondence among models at different levels or with different focus?	

Continued

I am a kind of	Scientific Reasoning This design pattern concerns a scientific problem to solve or investigate. Do they effectively plan ...
These are kinds of me	Interpretation of Dynamic Graphs Students are provided with a model in which they can manipulate model parameters and view the result...
I am a part of	
These are parts of me	Model-based reasoning
Educational standards	
Exemplar tasks	Force Concept Inventory (FCI) A multiple-choice measure of students' conceptual knowledge of forces, including Newton's 1st-3rd la...
Online resources	
References	diSessa, A. A. (1993). Toward and epistemology of physics. Cognition and Instruction, 10(2 & 3), 105-225. Hunt, E., & Minstrell, J. (1994). A cognitive approach to the teaching of physics. In K. McGilly (Ed.), Classroom lessons: Integrating cognitive theory and classroom practice (pp. 51-74). Cambridge, MA: MIT Press. NSES standards. Stewart, J., & Hafner, R. (1994). Research on Problem Solving: Genetics. In D. Gabel (Ed.), Handbook of Research on Science Teaching and Learning (pp 284-300). New York: MacMillan. White, B. Y., & Frederiksen, J. R. (1998). Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. Cognition and Instruction, 16(1), 3-118.