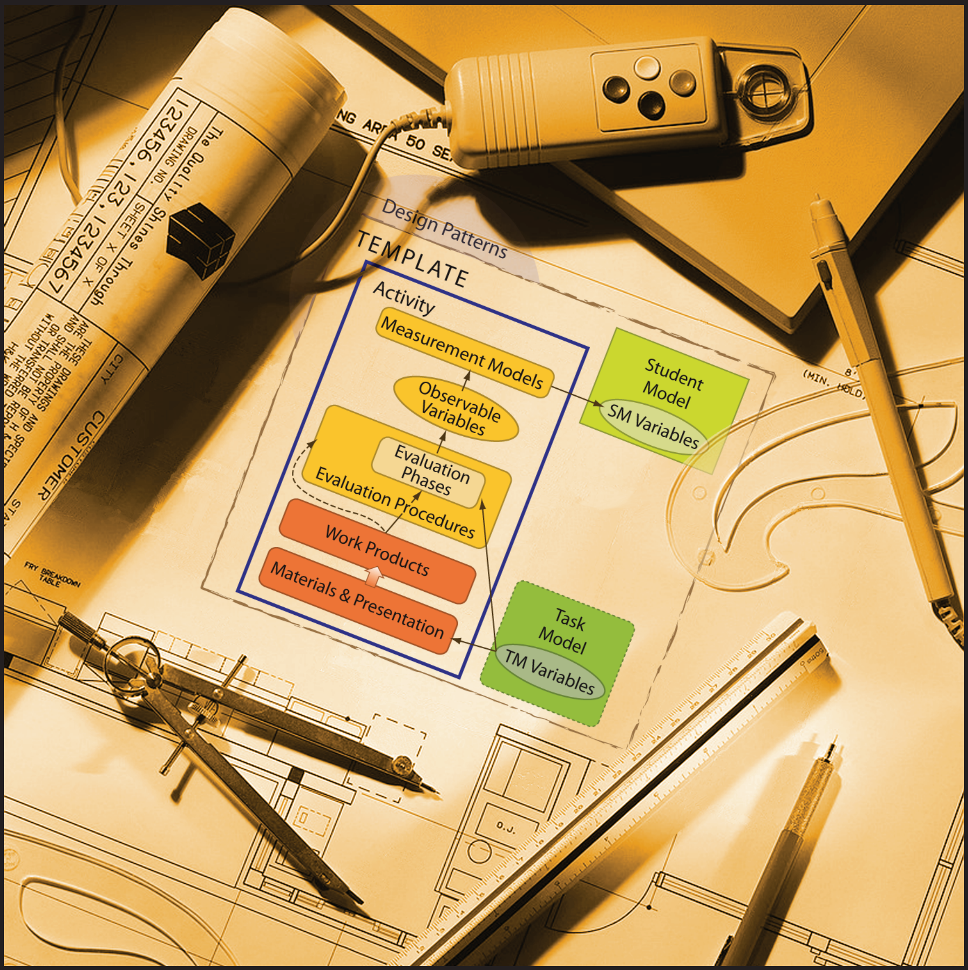

PADI Technical Report 22 (DRAFT)
March 2008

Design Patterns for Assessing Internal Knowledge Representations

PADI | Principled Assessment Designs for Inquiry

Kavita L. Seeratan, Pedagogical Science LLC/University of California, Berkeley
Robert J. Mislevy, University of Maryland

Report Series Published by SRI International

SRI International
Center for Technology in Learning
333 Ravenswood Avenue
Menlo Park, CA 94025-3493
650.859.2000
http://padi.sri.com

PADI Technical Report Series Editors
Alexis Mitman Colker, Ph.D., *Project Consultant*
Geneva D. Haertel, Ph.D., *Co-Principal Investigator*
Robert Mislevy, Ph.D., *Co-Principal Investigator*
Lynne Peck Theis and Ron Fried, *Documentation Designers*

DRAFT

Technical Report 22

DESIGN PATTERNS FOR ASSESSING INTERNAL KNOWLEDGE REPRESENTATIONS

by

Kavita L. Seeratan[1, 2] and Robert J. Mislevy[3]

[1] Pedagogical Science; [2]University of California, Berkeley
[3] University of Maryland, College Park

2009

TABLE OF CONTENTS

Insert Table of Tables and Figures here

Insert Abstract Here

1.0 Introduction: Assessing Knowledge Representations

Narrowly construed, in purely epistemological terms, knowledge representation refers to how informational content from the external realm is mentally organized, structured, conveyed, and made available for use when needed. As a fundamental purpose of formal education is helping students build useful representations of valued knowledge, educators need ways to monitor their efforts—in short, to assess students' knowledge representations. The purpose of this report is to sketch out a design space for tasks that can be used to assess aspects of knowledge representation. The substance for the proffered design space draws on research on knowledge representation from cognitive psychology, which will be reviewed in Section 2, and experience in that context for eliciting and characterizing knowledge representations, which will be reviewed in Section 3. The design space is organized around structures called *design patterns* (Mislevy, et al., 2003), in terms of the elements of assessment arguments: the knowledge or proficiencies of interest, other knowledge that may be required in tasks, characteristic and variable features of tasks, and ways of eliciting and evaluating evidence from students. Section 4 provides an overview of *design patterns*, and Section 5 presents and discusses five *design patterns* useful for assessing aspects of knowledge representation:

- Building knowledge representations
- Updating & correcting knowledge representations
- Using knowledge representations
- Recognizing and producing equivalence of meaning across representations
- Re-expressing information across representations

Section 6 concludes with a summary of key ideas.

1.1 Some Terminology

Although internal representation of the perceived external world, referred to as the 'knowledge that one holds about external reality,' is fundamental to the definition of knowledge representation, it is certainly not the whole matter. Knowledge representation and reasoning about knowledge are inextricably intertwined. One cannot talk about the first without also, unavoidably, discussing the second.

While it is true that the topic of 'representation' lies at the core of cognitive psychology, it is also true that the more frequently it is used, the more interpretations and operations we can expect. As Von Glasersfeld (1987) has pointed out, "…the word is used with a rather wide range of meanings" (p. 215), even amongst cognitive scientists. For instance, some authors use 'knowledge representation' as a synonym for 'internal representation' (Fodor, 1981). Others use it as a synonym for 'thinking, imagining, and visualizing' — or, in other words, as a general term for knowledge (Finke, 1989). Yet others see it within a specific theoretical context such as 'cognitive representation' (Palmer, 1978), 'imagistic representation' (Kaput, 1988), or 'symbolic representation' (DeLoache, 1989; Nelissen & Tomic, 1996). While we will discuss some of the various opinions and controversies about knowledge representation in Section 2, there is a broader level terminology controversy in the field, which we feel should be addressed at the onset — that is, the difference amongst terms such as 'internal representations,' 'external representations,' and 'external representations of internal knowledge representations.' This paper focuses primarily on internal representations, but at varied points, we discuss the relation among the three.

Internal representation: Internal representation refers to the way that information about the world is represented in our brains, and as such lies at the center of learning, interacting, and problem solving. Knowledge representation at the level of *implementation* is a vital and often controversial branch of cognitive science, and as we will see, research on this question provides us with useful tools for discussing knowledge at the level of *information*. For pedagogical purposes, it is this informational level that is of primary interest — the contained information, the modes of representation, and the nature of the associations — aside from details of implementation. We note however, that it is less useful to think of a person's knowledge as fixed and static structures, than as networks of associations that encompass propositions, symbols, and images and procedures for working with them, which are activated, assembled, and adapted for particular situations, to achieve particular goals. An expert constructs a mental model for a particular physical situation around principles such as equilibrium and Newton's laws, for example. In contrast, a novice may pull together a pastiche comprising surface features of the situation, less systematic "phenomenological primitives" (diSessa, 1983), and analogies to more familiar experiences whether or not they would conform to "expert" models (Collins & Gentner, 1987).

External representation: External knowledge representations (i.e., EKR), or inscriptions (Lehrer & Schauble, 2002), are "physical or conceptual structures that depict entities and relationships in some domain, in a way that can be shared among different individuals or by the same individual at different points in time. They are human interventions that overcome obstacles to human information processing with respect to limited working memory, faulty long-term memory, over time and in volume, coordinating the actions of many individuals, and idiosyncratic ways of thinking about some phenomenon of common interest" (Mislevy et al., 2007). Some ideas cannot be expressed, communicated to others, replicated, or improved by cycles of evaluation and revision, unless they are captured in inscriptions (Lehrer & Schauble, 2002). Examples include graphs, symbol systems, and maps. Because external representations embody important patterns within and across domains, learning to use them can promote analogous patterns in internal representation. In assessment, they are used to convey information to students, support their work, and structure outcomes.

External representations of internal knowledge representations: In the study of cognition, inscriptions have been developed for the purpose of depicting, simulating, and understanding aspects of human knowledge such as how we perceive, think about, and act on the world. Examples of external representations of internal knowledge include knowledge networks, schemas or frames, and production systems such as Newell's GOMS and SOAR, or Anderson's ACT. Such tools can be used to describe salient aspects of a person's internal representations without claiming isomorphism at the level of implementation. Rather, the contents, relationships, and procedures depicted in a representation may reflect salient features of a person's knowledge. Cognitive researchers use these representations as tools for studying knowledge representation, instructional designers use them to explicate targeted knowledge in curricular domains, and assessment designers use them to help create assessment tasks and evaluate students' performances.

1.2 The Importance of Assessing Knowledge Representations

Assessing how a student organizes knowledge can provide pertinent insights into the quality of knowledge thus far attained and suggest pathways to facilitate future knowledge acquisition. The connection, organization, relation, interrelation, and integration of knowledge, at general and specific levels and sublevels, are all important in deciphering not simply the amount of knowledge a student possesses, but the character of knowledge — specifically, whether it is organized for effective action and efficient learning. Knowledge representation encompasses structural and functional relations among concepts and connotes dynamic structures that allow the mental manipulation of ideas for problem-solving, inference, and/or predictions related to the knowledge system. The level of detail of the hierarchical knowledge network is a gauge of an individual's conceptual understanding; the intricacies can influence the quality of performance on cognitive and academic tasks.

Variation in the structure and function of expert and novice knowledge relates directly to the observed variation in performance and wide array of individual difference possible with regards to level of understanding of conceptual content. Research suggests that experts have not only acquired extensive knowledge but that they have organized that knowledge into dense, highly interconnected, hierarchical layers that are contextualized. Experts appear to possess an efficient organization of knowledge with meaningful relations among related elements, clustered into related units that are governed by underlying concepts and principles. This principle suggests that 'knowing more' means having 'more conceptual chunks' in memory, 'more relations or features' defining each chunk, 'more interrelations' among the chunks (Bransford et al., 1999). Compared to novices, experts appear to have knowledge representations that are more stable, integrated, and linked to purposive actions (Glaser, 1988).

Consequently, assessing a student's knowledge representations in a particular domain provides important information about his or her deep conceptual understanding. In this report, we will discuss issues surrounding the assessment of knowledge representations and the use of *design patterns* for improving our capacity and efficiency for doing so.

2.0 An Overview of Knowledge Representation

How is knowledge represented in the human mind? This question is not new but rather has been and continues to be at the forefront of research in the field of cognitive psychology, although not exclusive to this field. Over the years, many have attempted to provide their own model, definition, and explanations of knowledge representations and how they work. In the section that follows, we provide a brief synopsis of some of these views.

2.1 What is a Knowledge Representation?

Craik (1943) stated that if a learner were to carry a small-scaled model of his or her external reality within his or her head, then he or she could potentially make use of that model to analyze and tease apart its individual components, try out different alternative solutions and possible actions, reason about optimal solutions to problems, and utilize his or her knowledge and understanding of the world whenever needed in the future (St-Cyr, 2002). Generally speaking, knowledge representations are internal representations of external reality that enable an individual to make inferences, deductions, and predictions, to understand and explain a process, and to decide on the best course of action to take in particular situations or contexts (Johnson-Laird, 1983). Knowledge representations can be characterized as what we know about the real world and how it is represented and organized in our brains. The interest in this topic has driven researchers to try to operationalize and objectify the construct. However, like most things that are unobserved, the nature of the 'language of the mind' (or whether there is indeed such a thing) is a matter of debate even among cognitive scientists (see, for example, the debate between Greeno & Moore, 1993, and Vera & Simon, 1993, and the integrative response suggested by Clancey, 1993).

Johnson-Laird (1983) observes that "human beings understand the world by constructing working models of it in their minds" (p. 10). But what would an internal working model of reality entail? Deloache (1989), for instance, proposes that a cognitive or internal representation refers to an organized system of knowledge, which reflects certain, albeit not all, information about the reality that is being represented. Dretske (1986) suggests that a representational system is "any system whose function is to indicate, by its various states, how things stand with respect to some other object, condition, or magnitude" (p.102). Others, such as Kosslyn (1978) and Anderson (1990), distinguish between two kinds of internal representations: *iconic* and *propositional*. Iconic representations are comprised of two levels. A *level 1* iconic representation, characterized as 'photographic,' is, as its analogy with photography implies, closely related to direct perception, having the same structure as external reality (Nelissen & Tomic, 1996). A *level 2* iconic representation, in contrast, is influenced by perception but also by interpretation and is much more removed from reality than *level 1* and, hence, more closely tied to propositional knowledge (Kosslyn, 1978). According to Kosslyn, fundamental cognitive, non-perceptive processes (such as conceptual thinking) guide the construction of *level 2* iconic representations — for instance, clarifying relationships that cannot be perceived directly. *Propositional representations*, on the other hand, are comprised of symbols, structures, networks, and propositions, which Anderson termed meaning-based knowledge (1990). With propositional representations, structure is both constructed and interpretive.

Correspondence epistemology sees knowledge as a simple mapping or reflection of the external world, where every conceptual object (symbol) must correspond to one or more physical objects in the environment; it is a homomorphic map that encodes the structure of external reality (Heylighen, 2001). An alternative epistemology, constructivism, posits that knowledge is not a passive mapping of the outside world but rather an active construction on the part of the learner. This construction does not claim to reflect an objective reality but rather is supposed to aid the learner in adapting or fitting into a world that is subjectively experienced (Heylighen, 2001). In such an epistemology, knowledge is not mapped one–to–one with external reality but is coherent with other internal models of knowledge (Thagard, 1989). With this perspective, the lack of direct access to exterior reality no longer constitutes an obstacle. These two epistemologies (i.e., correspondence and constructivism) have distinctions parallel to those between the iconic and propositional representations discussed previously.

Mental models are central to theories in which individuals represent, interact with, and/or create the world through symbols (e.g., Knorr-Cetina, 1981; Mead, 1964). Yet another perspective, termed the connectionist viewpoint, posits that learning does not proceed by means of symbols but rather is based on recurring patterns — patterns that an analyst, from an external perspective, might refer to as schemas, scripts, or frames (e.g., Anderson & Bower, 1973; Carley & Palmquist, 1992; Fiske & Taylor, 1984; Tversky & Kahneman, 1980), in which the relationships among the elements provide the desired information. Connectionists have a network orientation to mental models. All units are built up out of elementary units and can be connected and activated in various ways. Mental models are externally represented as networks of concepts with specified relations and interrelations among all elements. Such networks can be analyzed into their components: concept; relationship; statement; and map (Carley, 1984; Carley & Palmquist, 1992). A *concept* (e.g., a penny) is a single idea bereft of deep meaning except when connected to other concepts (e.g., dime). A *relationship* is the connection that links two concepts together (e.g., is greater than). Relationships can vary according to directionality, strength, sign, and meaning. A *statement* allows us to specify two concepts and the relationship between them (e.g., a dime is greater in monetary value than a penny). And finally, a *map* is a network formed from one or more statements (e.g., a dime is greater than a penny and a dime is smaller than a quarter); both of these statements share the concept of dime. The resultant network or map is a representation of the initial beginnings of a mental model.

In line with a connectionist, or associative network view, Kintsch (1998), among others, proposes to represent knowledge as a network of propositions that he calls knowledge net. The nodes of the net are propositions, schemas, frames, scripts and so forth; the links are unlabelled and vary in strength. The meaning of a node is given by its location in the net — that is, by the strengths with which it is linked to its neighbors, those immediate and those steps apart. According to cognitive scientists with a connectionist framework, to be knowledgeable about some particular domain or aspect of a domain is to understand the structure, relations, and interrelations among the important concepts that define that domain. Being knowledgeable also requires changes (e.g., assimilations and accommodations) to the structure and organization of existing knowledge nets as further skills and expertise are acquired.

2.2 Definition of Knowledge Representations

Representation is seen as a mapping between two worlds, one representing the other. Markman's (1999) definition of knowledge representation, for example, has four components: a represented world, a representing world, representing rules, and a process that uses these representations. The terms *represented world* and *representing world* comes from a paper by Palmer (1978). The former refers to what is being represented (i.e., the referent) or the domain containing the objects that the representations are about, and the latter refers to the domain that contains the actual representations (i.e., the representational medium). *Representing rules* refer to the rules or relativity between the represented world and the representing world that map elements of one to elements of the other. Finally, there is the *process that uses the representation*, without which the representations themselves would be senseless. For amplification of Markman's (1999) definitions, consider Gallistel's (1990) proposal for the meaning of representation:

> "…a representation is a correspondence between the formal structure of the *represented system* and the formal structure of the *representing system* that enables one to predict results in the represented system on the basis of operations conducted within the representing system. (…). To say that the brain represents aspects of the world is to say that there is a correspondence between the formal structure of processes in the nervous system and the formal structure of the represented aspects of the world that enables operations in the nervous system to anticipate relations in the environment and to predict the consequences of acting in specified ways on or within that environment" (p. 582).

The main features of the Markman/Gallistel view of representation apply to both internal and external representations, in terms of represented and representing systems and support of operations. These elements are explicit and formal in the case of external representations, and inferred in the case of internal relationships. In particular, cognitive scientists' external representations of subjects' internal representations, in the form of production rule systems or knowledge networks, for example, are generally more organized, explicit, stable, and rule-based than the system they are meant to represent.

2.3 Characteristics of Knowledge Representations

Knowledge representations are characterized by certain properties of which those central to our purposes will be discussed. One of the most important is that knowledge representations are *not all inclusive* — that is, they do not attempt to include everything present in the represented world, rather only certain entities and relationships that are considered significant. Some of the limitless complexities of the real world must be omitted in the representations. There will, of course, be situations where what is omitted from the mapping is indeed crucial for the context of the real world situation at hand. Representations will have to be modified and adapted to include contextual factors if successful problem solving becomes hampered (Mislevy et al., 2007; Nelissen & Tomic, 1996).

Another important property of knowledge representations is that selecting a representation means making a set of *ontological commitments*. These commitments bring certain aspects of the real world into focus at the expense of blurring other aspects. This focusing effect is a critical aspect of knowledge representation, offering a simplification of the natural world because it can be cognitively overwhelming. The connectionist perspective described earlier, for instance, is only one way of thinking about the complexities of how knowledge is represented in the brain. This is indeed a useful perspective, adopted by many. But as our discussion also indicated, this is not the only possible perspective. This specific perspective brings certain principles into focus while blurring other principles; other principles could be highlighted in different models of knowledge representation (Mislevy et al., 2007; Nelissen & Tomic, 1996).

A third property of knowledge representations is that their *attunement to uses* accounts in part for the multiple internal representations possible in a domain (Ainsworth, 1999; Mislevy et al., 2007). When the complexities of the external world lend themselves to being represented at different levels or from different perspectives, one can have multiple internal representations, with each knowledge representation highlighting aspects, entities, and relationships at different levels of analysis. For instance, a tabular display of target mathematical functions can demonstrate the same reality or meaning as a graphical display. Being able to view and integrate knowledge about external reality from different perspectives facilitates deeper knowledge and understanding about the represented world and is a sign of developing expertise (Seeratan, 2006[1]).

Yet another property of knowledge representations is that they act as *a medium for efficient computation*. Representations offer ideas about how to organize information in ways that facilitate making inferences. Again, different representations can be more efficient than others for different purposes. Knowledge representations are also *a medium of human expression*, providing a common language for individuals to express information in ways that incorporate experience from other times and other people.

Other characteristics of knowledge representations, as stated by Ok-Choon and Stuarts (1995) include: incompleteness, limited stability, no firm boundaries, parsimoniousness, abstractness and generality, and image-likeness. For further discussion of these characteristics, some of which overlap with those already discussed here, please refer to the article by Ok-Choon and Stuarts (1995).

2.4 Functions of Knowledge Representations

Knowledge representations have many important functions other than providing an internal representation of external reality. The first function is an *organizing* function — in particular, organizing reality and cognitive processes. Through organization, knowledge is gained about, and control is gained over, reality. Even more substantial, through organization we gain control over our own cognitive functioning, as cognitive processes are directed by our internal representations of knowledge (Nelissen & Tomic, 1996).

Another important function of knowledge representations, also at the level of organization, is to layer *knowledge at succeeding levels of higher order thinking* (Nelissen & Tomic, 1996). Transitioning meaning from one representational system to another (e.g., from equation to graph in mathematics) creates a deeper understanding of the concept at hand. Multiple representations of *meaning* play an important role in learning to the extent that they facilitate the construction of coherent mental models, which in turn are thought to reflect higher levels of cognitive functioning (Dienes, 1973; Kaput, 1989; Moss & Case, 1999; Seeratan, 2006[1]). According to Kaput (1989):

> “cognitive linking of representations creates a whole that is more than the sum of its parts. . . . It enables us to 'see' complex ideas in a new way and apply them more effectively" (p. 179).

Learners will construct references across different representations if they are provided with a rich source of domain representations. Such knowledge then can be used to expose the underlying structure of the domain represented — that is, abstracting the common concept preserved within the various representations. Dienes (1973) argued that having the same concepts represented in varying ways provides learners with opportunities for building abstractions; learners come to discover invariant properties of a domain in the face of perceptually salient, but conceptually irrelevant, differences in the appearance of multiple representations (Ainsworth, 1999; Nelissen & Tomic, 1996; Seeratan, 2006[1]).

Another function of internal knowledge representations is a *communicative function*. Vygotsky (1977) believes that as a result of interactions with adults, children develop language. Internal representations regulate their interaction with adults and subsequently, the child’s own behaviors. A final function of knowledge representations to be mentioned here is its *psychological function*. According to Nelissen and Tomic (1996), an individual must learn a number of important things: a) that a representation of reality is not itself reality, b) that his or her representation of the world can differ from that of another person, c) that his or her representation can be subjected to reconstruction, d) representations are only approximations of reality, e) misrepresentations of reality may be possible, f) in many cases, only a few dimensions of reality’s complexity can be represented, and g) representations may be consciously or unconsciously constructed.

2.5 Initial Processes with Knowledge Representations

Often used in the domain of physics, the black-box analogy is a tool employed for helping to conceptualize or explain the idea of an internal knowledge representation (Rouse & Morris, 1986; St-Cyr, 2002). The logic behind it is as follows: an individual is given an input and an output (i.e., something goes into the black-box and something else comes out from it). The task is to decipher how the input gets transformed within the mind to produce the desired output. Here, the black-box is being used as an analogy for the mental model. The black-box includes one's internal representation of the external reality (i.e., the input) as well as its structures, connections, interconnections, and relationships to concepts and sub-concepts (resulting from prior knowledge and experience), which are related to the input. The individual must use knowledge they hold about the system to establish mappings and relationships between the input and output; they must also use this knowledge to make predictions, adaptations, reasoning, and for problem solving. Hence, if an individual can explain the contents of the black-box, then they can be presumed to possess an internal model of the system. If they can explain the contents accurately and in much detail, then they are deemed to have a good level of understanding as opposed to if they can explain the contents only somewhat (Rouse & Morris, 1986).

To further elucidate the process of building knowledge representations, Kintsch (1996) makes some useful observations about the *theory of direct perception*, of which Gibson (1977) was a principal advocate. Gibson — via the theory of direct perception — claimed that individuals do not need mental representations of the world in order to perceive it. Gibson suggests that percepts are not mediated by cognitive processing in the brain but rather that information from the external world is picked up directly from the environment as it is. No inferences or top down processing is required for perception simply because the input stimulus is deemed to contain all information necessary for visual perception. Gibson also rejected the theory that meaning is provided through long-term memory via associations to prior experience and knowledge (i.e., indirect perception or thinking about the system at a higher level). Rather, he argued that the potential uses of objects are directly perceivable, as affordances that can be picked up by the visual system (St-Cyr, 2002).

Kintsch (1996) asserts that although Gibson's theory of direct perception may be valid, it does not account for the knowledge one holds about the world. Kintsch takes an *indirect theory view*, suggesting that an individual might indeed be able to perceive a system and all its components, with respect to affordances and invariants when it exists in the external world, and hence be ready for interaction. However, before doing so, he or she will need to stop and think about all the different functions of the system, analyze its structure, relations among components, and understand its proper functioning (i.e., enrich, organize, and make sense of the data obtained from the visual system). Hence, the internal representation of information includes both perceiving the system at a lower level (i.e., Gibson's direct perception) and thinking about the system at a higher level of abstraction (i.e., indirect perception). Furthermore, as Kintsch (1996) points out:

> "it is important to know that higher levels of representation do not replace the lower levels, but encapsulate them, resulting in a complex and not easily analyzed system of representations" (p.32).

Another related question that surfaces when considering the construction of knowledge representation concerns the representation code of the stimulus to be perceived. Information in the external environment may exist in different ways. For instance, an individual may have to read text, view an image, or maybe listen to sounds. Are there different cognitive subsystems for each type of information code? According to Paivio (1986), who focused specifically on verbal and image–oriented information, this is indeed the case. Words and sentences are thought to be encoded and processed only within a verbal system, whereas images are thought to be encoded and processed within an image system and a verbal system (Schnotz & Bannert, 2003). Following Paivio's theorizing, Mayer (1997) postulated that because text and images are processed in different cognitive subsystems, each system constructs its own mental model. Mayer suggested that verbal selection processes generate a text–based mental model; similarly, image selection processes generate an image–based mental model. Both mental models are then integrated through a one–to–one mapping process — that is, elements of the text–based model are mapped onto elements of the image–based model, and vice versa. Similarly, relations within the text–based model are mapped onto relations within the image–based model, and vice versa (Schnotz & Bannert, 2003). The congruence between text processing and image processing, assumed in Mayer's (1997) extension of Paivio's model, seems questionable. There are sharp modality–specific distinctions between the processes used for verbal information processing and codes used for visual information processing. Text and image are different sign systems, which may influence the way information is encoded and processed and consequently result in fundamentally different forms of *mental representations* (Seeratan, 2006[1]). Unlike mental representations, however, mental models are not confined to specific sensory modalities. For instance, a mental model of a particular concept can be constructed not only by visual perception but also by auditory, kinesthetic, or haptic perception. A mental model may contain less information than the corresponding visual image or text because details are necessarily omitted. Simultaneously, a mental model may contain more information than the visual image or text because it includes inferences based on the learner's prior knowledge, which is not present in mere perception (Kintsch, 1998; Schnotz & Bannert, 2003). Hence, if differences exist in the way meaning is extracted from text as opposed to how meaning is extracted from images, it is conjectured they would be at a level preceding the construction of the mental model (Seeratan, 2006[1]).

2.5.1 Organizing and Constructing Knowledge Representations

Evidence suggests that memory plays an important role in forming and retaining mental models (Garnham, 1997). Dutke (1996) suggests that individuals construct mental models based on patterned information already available in long-term memory, referred to here as schemata. For example, an educational psychologist must have general knowledge about psychology, education, human development, and even quantitative and qualitative methods. This knowledge, gained through education, training, and experience, is stored in long–term memory. When facing an issue, the psychologist will go through a problem solving situation, try to predict outcomes, and find a solution. She will gather together knowledge available and embedded in long–term memory and dynamically construct the most accurate and complete mental model possible at the time that will allow her to achieve efficient decision–making. Note that there is some risk of reification here. We must bear in mind that the inscription based on behaviors elicited from a subject and called a schema by a researcher is a new and unique construction, partly of the subject as the behaviors arose from patterns of associations available to the subject and assembled to meet the demands of the elicitation procedure, and partly by the researcher through the choice of methods for elicitation, analysis, and representation.

Understanding the structure of expert knowledge provides invaluable insights about the nature of thinking and problem solving. It also provides information about the conditions necessary for achieving or pursuing optimal knowledge. Research suggests that experts have not only acquired extensive knowledge but that they have organized that knowledge into dense, highly interconnected, hierarchical layers that are contextualized. The concept–mapping tool provides a graphical representation of how knowledge is presumed to be represented and organized in the brain of experts. Figure 1 provides an illustration of a concept map where the concept of focus is paradoxically a concept map. It is what an expert's knowledge web might look like for the concept of "concept map." This concept map illustrates the intricate network of hierarchically connected, organized ideas, relations, and interrelations among various levels of the concepts and sub–concepts (i.e., super–ordinate and sub–ordinate levels). Each fact is organized into a relevant category and correspondingly linked with other related facts, sets of facts, or categories of knowledge; common features of categories/ideas/relationships explicit (Seeratan, 2006[2]).

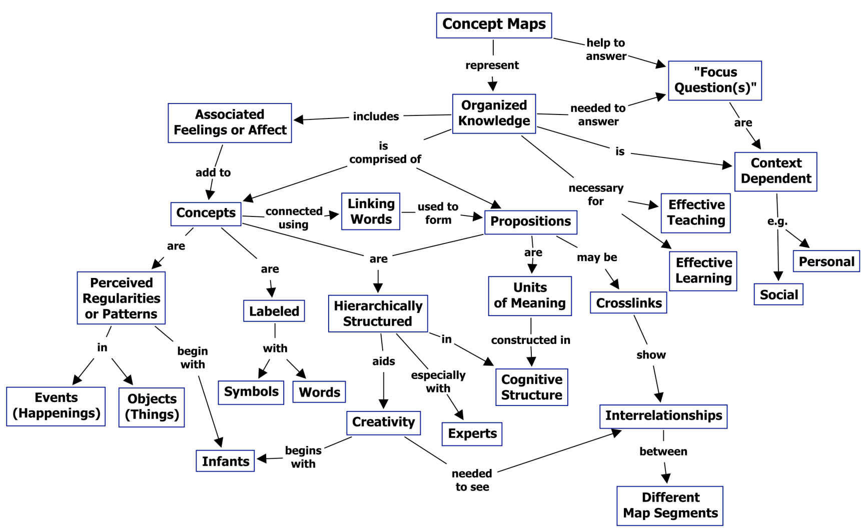

Figure 1. Graphical Representation of a Concept

Because the world is filled with constant flow of information, the brain must find ways of organizing and indexing information for later retrieval and problem solving. Research tells us that there are indeed limits on how much information can be stored in short–term memory, theory stating that it can hold roughly 7 +/- 2 items or pieces of information. Being able to organize and chunk (i.e., group similar items) information into patterns of meaningful categories enhances one's cognitive processing, memory, and problem solving capacities (e.g., Glaser & Chi, 1988). Hence, in knowledge representation construction, an individual is faced with complex input from the external world. Based on prior information and experience stored in long–term memory in the form of schemas, the individual tries to establish correspondences between what is externally presented and what is already existing internally (i.e., internal models). If an individual has well developed schemata (e.g., expert), their perception can be optimally guided to integrate new information into existing mental maps or create new maps. They quickly will be able to identify items or information that fit within their existing schemas (Bransford, 1999). They also will be able to identify when the new information is not in accordance with what is already there, recognizing the need to make adjustments or changes to their existing schemata or creating new ones (to be discussed further in a later section). The important point here is that the more intricate the schemata, the more efficient the process of knowledge construction.

Knowledge representations are important factors that help determine human performance on cognitive tasks. Given the same set of stimuli or concepts, different people may perceive or focus on different interrelationships or aspects among these concepts and thus develop different structures to organize and represent the concepts. For instance, those with extensive prior knowledge or experience about a particular topic will use these informed lenses in perceiving and eventually analyzing new content. On the other hand, those who have not yet formed related schemas will tend to have much less direction and focus in perception. If an individual has little existing knowledge or poorly organized preexisting knowledge, then they may be lead in the direction of trying to memorize facts or forming categorizations that are based on superficial features of the task as opposed to deep principles (Chi et al., 1981).

The task of organizing knowledge around core concepts, principles, or big ideas will be more difficult, if not impossible, without an organizational blueprint to guide the process. If knowledge is stored as numerous isolated facts, unattached to other parts of the knowledge system, the information will not be used and may not even be retrievable in situations where it is relevant. Different knowledge representations lead to different ways of accessing and retrieving these concepts when performing a cognitive task — which in turn, can cause performance differences on the task. Knowing more means having ”more conceptual chunks” in memory, ”more relations or features” defining each chunk, ”more interrelations” among the chunks (Bransford, 1999). Experts not only have acquired knowledge but are also good at retrieving the knowledge that is relevant to a particular task; their knowledge includes a specification of the contexts in which it is useful (Bransford, 1999; Larkin, et al., 1980; Glaser, 1992). From the perspective of Walter Kintsch's Construction–Integration (CI) theory of comprehension, as described in the following section, relevant and useful associations are likely to be activated, and once activated, provide connections through higher-level schemas for activating procedures, alternative representations, and problem–solving strategies.

2.5.2 Knowledge Representation and Comprehension

Comprehension can be thought of as constructing a representation of a particular and unique real–world situation in terms of the schemas and representations a person brings to that situation — that is, building a situated knowledge representation to suit the demands of the occasion. Walter Kintsch's (1988, 1998) Construction–Integration (CI) theory of comprehension offers many insights into the role and importance of knowledge representation in this process.

Although Kintsch originally proposed the CI theory for reading comprehension (notably Kintsch & Van Dijk, 1978) and still focuses attention on comprehension of written texts and propositional representations, he notes that the theory applies more broadly to other forms of representation, both internal and external, such as images and "runnable" mental models, and has applied the approach with scientific and arithmetic models (e.g., Kintsch & Greeno, 1985). In the following discussion, therefore, "text" refers to the stimuli being perceived and not necessarily to linguistic content; and "situation model" refers to internal representations involving meanings and patterns beyond propositional content.

According to Kintsch, comprehension proceeds through two representational phases: text–based and situation–based representations. Text–based representations, believed to be the first step in the process of comprehension, are "those elements or links in the mental representation that have a direct correspondence to elements in the text," as in verbatim recitations, simple paraphrasing, or summarizing (Kintsch, 1998). The situation–based model, created next, is established in working memory through a process of integration and expansion whereby information that has been extracted from the "text" combines with prior knowledge and experience from long–term memory (Gernsbacher, Varner, & Faust, 1990; Kintsch, 1998; Shafrir, 2001; Seeratan, 2006[1, 2]). This procedure allows new information to be represented, understood, and subsequently added to the cached mental models, allowing them to become highly structured and more elaborated (Johnson-Laird, 1983; Tversky, 1997; Van Dijk & Kintsch, 1983). The two representational phases of meaning construction exemplify two often confused knowledge processes — mental representation formation (which corresponds to the text–-based representation) and mental model formation (which corresponds to the situation–based representation). A mental representation is an internal depiction of an external situation or concept with minimal interpretation, linking, or elaboration with existing knowledge (text–based). A mental model, on the other hand, refers to an explanation in someone's thought process for how something works in the real world. It is an internal symbol of external reality that has been linked with prior knowledge and is fundamental to comprehension and interpretation of meaning (Seeratan, 2006[1]).

Once knowledge representations are built, organized, and constructed, it is important for learners to understand the mental model they have constructed. If users hold a mental model of a system but do not understand it, they will likely not use it properly to correctly solve problems or make predictions or inferences. Rasmussen (1985) presented a framework called abstraction hierarchy that aims to explain the functioning and malfunctioning of complex systems. This framework, also described by St-Cyr (2002), is comprised of means–ends relationships which provide answers to the questions such as what, why, and how with respect to a component, function, or relation present in the system. Rasmussen suggests that this framework can be used to enhance comprehension of mental models. Please refer to the Rasmussen (1985) article for further details.

2.5.3 Elaborating and Revising Knowledge Representations

Constructed dynamically, mental model organization is highly dependent on the capacity of the learner to update, elaborate, and correct his or her mental model as necessary. Updating and revisions occur when coherence between new information and old information is broken, when there are changes to structural relations, and/or when new contradictory information is presented (Albrecht & O'Brien, 1993). For instance, consider the following simple example: the red car is *old* and does *not function properly*. This statement highlights two local defining characteristics of the red car: that it is old and does not work well. Now consider if another statement is given: the red car was driven across country without mechanical issues. This statement is a global statement about the red car that is incongruent with the local representations. From this example, we see that coherence between local and global representations has been broken, and this incongruence signals to the learner that something in his or her mental model is incorrect, incomplete, and needs to be updated.

2.5.4 Using Knowledge Representations

Moray (1999) suggests that mental models help users *reason about complex systems* and thus *facilitate the learning process*. Consider the example of an individual who worked with a system for a long extended period of time and solved, let's say, over 100 different problems. Over this time period he or she developed a mental model that clearly delineated concepts and interrelations among concepts in an intricate hierarchical format. If a new situation presents itself, the individual will need to construct and use another mental model for solving it. When constructing this mental model, generic knowledge as well as previous mental models will be available for modifications or extensions. If there is enough overlap between past models and the current situation, then new elements may be assimilated or integrated into existing models; if common elements are few, then necessary accommodations will be made. Hence, the use of previous models or parts of models enhances the learning process (St-Cyr, 2002). Kieras and Bovair (1984) found that learning how to operate a device is affected by understanding a model that describes the internal mechanisms of the device. In fact, having a mental model of the device *contributed to learning*, ensuring that the characteristics, operations, and relationships within the system were remembered. Another study by Dixon (1991) demonstrated a related finding — that knowledge representation about one device can *help the user to learn a new device* if the two devices have some similar operating procedures (St-Cyr, 2002). In order for this to be true, the knowledge representation must be accurate and detailed.

A final use of mental models to be described here is that they can help to *explain knowledge and skill acquisition*. As this paper indicates, people use existing mental models to organize new knowledge in meaningful ways and to develop a set of skills to help solve problems. They also use the knowledge they acquire to construct more accurate and complete mental models. Having more accurate mental models that are hierarchically based will support further knowledge and skill acquisition in turn.

2.6 Relation of Knowledge Representations to the Classroom

Knowledge representations can guide subsequent training/remediation and assessment for skill/knowledge acquisition. Prior experience and knowledge — in particular, the organization of this information into schemata — influence future learning and mental model formation. The assessment of differences between the knowledge representation systems of experts and novices is useful in guiding the training of novices for skill acquisition and in assisting the design of tasks and tools for performance and understanding enhancement. For instance, curriculum or training material can be designed to target differences between novice and experts in knowledge representation, highlighting key characteristics of expert knowledge representation for facilitating novice learning. The presentation of concepts should be consistent with expert's structural representation of concepts. Content should be taught (and assessments administered) in ways that facilitate accurate and detailed representation of knowledge in organized hierarchical ways that will in turn facilitate not only learning acquisition but also the retrieval process.

2.7 Knowledge Representations and Assessment

We find knowledge representation throughout the enterprise of educational assessment. An assessment is in itself an external knowledge representation; it makes explicit, sharable, and public the knowledge that is valued, the ways it is used, and the standards of good work. The analysis of any domain in which learning is to be assessed must include identifying the knowledge representations that people work with and how they use them. Because developing facility with knowledge representations is central to developing expertise in any domain, claims about students' capabilities will be central to assessment: capabilities to choose representations, express information and obtain information from them, transform information from one representation to another, and so on.

Assessment tasks can be structured around the knowledge, relationships, and uses of the domain representations. Every assessment task must use forms of knowledge representations to provide information, offer affordances, and capture actions. Explicitly designing tasks around the forms people use and the ways they use them is key to assessing aspects of proficiency with them (Gitomer & Steinberg, 1999). This includes choosing efficacious representations for a given purpose, mapping situations into representations, reasoning through representations, translating information from one representation to another, and recognizing when to do so.

The discipline of assessment employs knowledge representations itself. We use Toulmin diagrams to structure arguments, job analysis questionnaires to determine valued knowledge and skills, scope-and-sequence charts to map curricula, and test specifications to assemble tasks into assessments. *Task templates* organize the results of *domain analyses* into schemas that guide task creation, from Hively, Patterson, and Page's (1968) arithmetic item forms to open-ended science inquiry investigations (SRI, 2002). The *design patterns* presented in Section 5 are themselves representational forms, offered in the hope that assessment designers will find them useful in constructing tasks to assess aspects of students' knowledge representations.

3.0 Methods for Assessing Knowledge Representations from Cognitive Research

There are vast arrays of assessment methods used for extracting, eliciting, characterizing, and in some cases, measuring knowledge representations. Although most were initially developed in the setting of cognitive research, they provide insights for constructing educational assessment tasks, especially with regard to task features, work productions, and evaluation procedures. Based primarily upon a thorough analysis and decomposition proposed by Cooke (1994), a brief overview of these methods is presented below. We highlight some of the knowledge elicitation techniques available, how they differ from each other, some limitations and advantages, and when one should be preferred over another.

3.1 Observations, Interviews, and Task Analyses

External signs of the internal 'mental' models can be inferred by actively observing, without cognitive interruption, the learner as he or she performs a domain–related task or solves a domain–related problem (Cooke, 1994; Welbank, 1990). Observations allow the observer to record features and actions of the learner's behavior in naturally occurring settings; identify problem solving strategies that are not consciously accessible; and detect anomalies, errors, or unusual actions in the operant's behavior. Cooke (1994) presents three variations of the traditional observations method: a) *active participation* in which the observer takes a more active role in the observation process by asking questions or requesting that the learner carry out specific tasks to be observed; b) *focused observation* in which the observer focuses his or her attention on particular aspects of a detailed task as it is being carried out by the learner; and c) *structured observation* in which there is a prearranged protocol established with regards to features of importance for detailed observation.

The main advantage of observational knowledge elicitation techniques is that they minimally interfere with the learner's task, environment, and mental processes. A disadvantage is the potential impact that the presence of an observing experimenter — or "outsider" — could have on the dynamics of the situation being observed (Cooke, 1994; Cordingley, 1989). A more important disadvantage of these observational techniques, however, concerns the interpretation of data collected from observations — that is, how should the researcher express what is learned about subjects' knowledge representations? What aspects are important to include? Can information from various subjects be represented in a common form? In educational assessment, when inferences about students are framed in a given theoretical framework, the challenges are to structure performance situations to elicit relevant behaviors and have evaluation procedures to recognize and evaluate their key features. Some tasks used in physics education research, for example, are specifically designed to provoke common misconceptions (e.g., Hestenes, Wells, & Swackhamer, 1992).

Table 1. Examples of Knowledge Elicitation Techniques

Techniques	**Variations**
Observations	Active Participation
	Focused Observation
	Structured Observation
Interviews	
Unstructured	
Structured	Focused Discussion
	Questionnaires
Task Analyses	Functional Flow Analysis
	Cognitive Task Analysis

Interviews, another means of assessing student knowledge, are the most frequently used of all knowledge elicitation techniques (Cooke, 1994). Retrospective in nature, interviews can be direct or indirect and can contain questions that are explicit or implied. Interview methods can be divided into two major classes: unstructured or structured. Unstructured interviews are free-form — that is, neither the content nor the sequencing of topics is predetermined (Welbank, 1990). Rather, with unstructured interviews, answers to previous questions form the basis for further paths of exploration. Structured interviews, on the other hand, follow a predetermined format in which the experimenter asks specific questions and focuses on particular points of interest. These are more readily adapted to educational assessment. Cooke (1994) presents a range of structured interviews techniques including: focused discussion and questionnaires. Focused discussions center on specific types of information, including cases, goals, or diagrams.

Discussing a case study, for instance, converges on specific experiences, problems, and solutions. Forward scenario simulation, another focused discussion technique, uses simulation principles to focus on a case; specifically, an experimenter provides the participant with an initial situation (i.e., data), and the participant must use his or her problem-solving skills derived from his or her mental model of relevant knowledge to solve problems, simulate the system mentally, and predict future states. ACT-based representation of knowledge, inspired by Anderson's ACT model (Anderson, 1983), is a focused discussion technique that focuses on goals. It is designed to elicit a network of static knowledge about the domain and a set of procedures that are performed on that knowledge. Here, learners are asked to break goals into specific sub–goals or actions and to subsequently generate a set of production rules for representing the goal versus sub–goal and goal–versus–action relations. Questionnaires, the final form of structured interview to be discussed here, also can be used to obtain learners' specific thoughts about their knowledge base or components thereof. This type of questionnaire is not the typical survey prototype but rather one that includes open–ended questions about concepts, values, attributes, and relations. For a more comprehensive description of these questionnaires as well as more observational or interview knowledge extraction methods, please refer to Cooke (1994).

Task analyses, another knowledge elicitation technique, involve both observation and interviews with a focus on a specific task or function. In task analyses, specific tasks or functions are decomposed into their corresponding subtasks in order to identify behavioral implications, the functions of the expert, and the relationship of particular tasks to the job as a whole (Cooke, 1994). Examples include functional flow analysis and cognitive task analysis. Functional flow analysis involves the identification, description, and diagramming of all functions and sub–functions of a system; cognitive task analysis focuses more intensely on the cognitive processes or demands of the task. PARI is an example of a cognitive tasks analysis methodology (Hall et al., 1994).

3.2 Process Tracing

Process tracing is conducted to track a “real time” course of action while the learner is engaged in a specific task. The data collected are recorded according to a pre–specified type of medium (e.g., verbal report, eye movements, actions, etc.) and are subsequently analyzed to uncover cognitive processes or knowledge structures underlying task performance (Cooke, 1994). Examples of process tracing techniques include verbal reports and non–verbal reports. Other process tracing techniques include protocol analysis and decision analysis (see Cooke, 1994, for detailed descriptions).

Verbal reports, normally recorded to allow the experimenter to go back and conduct detailed observations and analysis, differ from interviews in that they occur concurrently and are tracked while the participant is carrying out the task. Underlying task–specific cognitive processes and the representations presumed to support them are subsequently inferred on the basis of the knowledge that the expert is able to articulate while problem solving. Despite their many benefits and popular usage, there are a number of important limitations associated with verbal reports as knowledge elicitation techniques. For instance, such reports are difficult and sometimes impossible to collect if: a) the task that is being performed already involves and requires verbal communication from the respondent (Bainbridge, 1979); b) the task has a high cognitive load requirement, making it difficult for the respondent to think, perform, and talk aloud at once; c) the involved cognitive mechanisms are beyond the plane of consciousness (i.e., automated) and, hence, not accessible for articulation; and d) the respondent lacks the ability or experience required to verbalize their knowledge, thoughts, or feelings about the task (Cooke, 1994).

Table 2. Type and Latency of Verbal Techniques

Verbal Techniques	**Latency**
Self Reports	Concurrent
Shadowing	Concurrent
Talk & Think Alouds	Concurrent
Critical Retrospective	Retrospective
Interruption Analysis	Retrospective
Group Discussion	Retrospective

As shown in Table 2, verbal techniques can be classified in a number of ways, including: a) *self-reports*, where the learner or respondent describes to the experimenter what he or she is currently doing to solve a problem (Cooke, 1994; Ericsson & Simon, 1984); b) *shadowing*, where a second respondent, normally an expert, describes the steps taken by the learner or first respondent to solve a problem (Clarke, 1987); and c) *talk aloud and think aloud* protocols, where the respondent is asked to say out loud everything he or she would say to himself or herself (i.e., his or her internal conversations) when solving a problem (Sasse, 1991). Other variants of the verbal technique include: d) critical retrospective; e) interruption analysis; and f) group discussion. These latter three differ from the other verbal techniques in that they are conducted retrospectively (i.e., after the task is complete). In *critical retrospective*, the learner is required to comment and report on the problem-solving steps and cognitive processes they used throughout the task. In the *group discussion* technique, both the learner and several experts offer verbal reports on the task completed. Finally, while the *interruption analysis* technique proceeds without requiring the implementation of a think aloud protocol during task completion, the learner may be interrupted for clarifications or verbal reports at points during the task when the experimenter has difficulties gauging the learner's knowledge processes or thoughts (Olson & Rueter, 1987). The concurrent verbal techniques are often criticized as a potential interference with the learner's thinking and task performance; retrospective verbal techniques, on the other hand, lessen the problem of interference but are diminished by memory limitations (Cooke, 1994).

Verbal reports are the most popular process tracing technique, but there are non–verbal techniques as well (Sackett, 1977; Cooke, 1994). *Non-verbal techniques* are used most often when data that must be collected cannot be easily verbalized (e.g., if the task requires high levels of concentration, has high cognitive demands, is automated, or is meta-conscious). Tracking a learner's eye movements as they complete the specified task is one example of a non–verbal report; it can provide invaluable information regarding important environmental cues, the learner's focal point, and/or specific search patterns. Other actions monitored and carried out during the task can also provide useful indications of the underlying cognitive processes. For instance, keystrokes, mouse movements, and mouse clicks on a computer can be recorded and used to provide data with high levels of precision, although the results may be difficult to interpret (Cooke, 1994; Kowalski & Van Lehn, 1988).

3.3 Conceptual Techniques

The last set of knowledge elicitation techniques to be discussed here is called conceptual techniques (Cooke, 1994). These techniques are useful for revealing one's understanding of domain concepts and the relations/interrelations between their structures and functions.

Conceptual techniques are often indirect methods requiring less introspection and verbalization than interviews and verbal report techniques (Cooke, 1994). These techniques most often are used to assess mental models and knowledge representations because they expose system structure and function and produce inscriptions that can be easily quantified. (In the terminology of assessment, they can be used to elicit key aspects of students' knowledge representations and yield work products that are straightforward to evaluate.) Three sets of such techniques that utilize a structured interview approach will be briefly described: concept elicitation methods, data collection methods, and structural analysis methods.

As Cooke (1994) suggests, almost all conceptual techniques begin with a set of concepts, whether they are objects, elements, or parts that are central to an understanding of the domain in question. *Concept elicitation* techniques allow the experimenter to collect data on the learner's comprehension of a set of concepts presented in a particular domain. Some concept elicitation techniques focus specifically on domain terminology while others focus on the learner's ability to relate, integrate, and group concepts together. The *concept listing* technique, for instance, asks learners to list critical domain concepts; the *step listing* technique is more specific, requiring learners to decompose, analyze, and list the steps involved in performing a domain–specific task; the *chapter listing* technique requires learners to go through a list of all or selected domain concepts and subsequently divide and categorize them into headings and subheadings as if to create a table of contents of a book in that domain. Such techniques reveal the learner's structural understanding within and between concepts (and sub–concepts) of a domain. Another concept elicitation technique is referred to as the *laddering approach* (Cooke, 1994; Shadbolt & Burton, 1990). The purpose of this technique is to uncover domain superordinate and subordinate relations between concepts through answers to specific questions or prompts. Having the learner ask himself or herself "why" questions, for instance, can lead to the revelation of superordinate connections. When the learner asks himself or herself "how" and "what" questions, subordinate connections will be revealed (Boose, 1986). This exercise will help create a graph of superordinate and subordinate relations between concepts, resulting in a hierarchical taxonomy of domain concepts. A number of computerized tools have been developed to elicit and analyze these data in the form of proximity networks (e.g., Schvaneveldt, 1990).

Another set of conceptual techniques is *knowledge proximity measures*. These techniques (also referred to, somewhat unhelpfully, as *data collection*) are primarily concerned with the degree of relatedness or the proximity between two or more domain concepts (Cooke, 1994). There are two types of data collection techniques: rating/ranking and sorting. *Rating and ranking* measures are considered simultaneously since their mechanisms for acquiring proximity estimates between pairs of concepts are similar. Here, participants are given pairs of concepts, one pair at a time, and asked to rate the 'relatedness' of each pair using a rating scale. Although a great tool for revealing the relationships and interrelationships between concepts and sub-concepts, this technique can become time consuming, especially when the number of concepts exceeds 25. Consider, for instance, that in a one dimensional context, 25 concepts would generate 300 pairs to be ranked. In the *sorting* technique, participants are asked to sort concepts into different stacks based on their degree of relatedness (Cooke, 1994; Geiwitz et al., 1990). Although sorting exercises can take a variety of forms, one of the most popular is the *card-sorting task* (e.g., Chi, Feltovich, & Glaser, 1981). In this task, participants are asked to sort cards depicting various concepts into several categories and subcategories in accordance with their understanding of the relations and interrelations between them. Using this technique in the domain of physics, Chi and colleagues (1981) found that novices produced stacks based on surface features such as the presence of pulleys or inclined planes, while experts produced stacks in accordance with physical principles such as equilibrium and Newton's Third Law. In general, the sorting techniques are less time consuming than rating or ranking techniques, but the tradeoff is decreased sensitivity to proximity differences (Cooke, 1994).

The final set of conceptual techniques to be discussed is the *structural analysis* wherein a descriptive multivariate statistical technique aims to reduce relatedness estimates for a set of concepts to a simpler form (Cooke, 1994). *Network scaling*, one example of a structural analysis technique, involves the generation of a graphical representation underlying the hierarchical structural and functional relations between different concepts of a domain. This graphical representation, generated by a pathfinder algorithm (Schvaneveldt, 1990), can be analyzed according to proximity estimates between concepts and network similarities between different possible answers. There are many other structural analysis techniques including *multidimensional scaling, clustering routines,* and *free association.* See Cooke (1994) for a more detailed description of these and other knowledge elicitation techniques.

These are all good ways for assessing and eliciting aspects of knowledge representations. As Cooke (1994) points out, these techniques differ in many ways, some being tradeoffs of others. Our goal in this paper is to make the reasoning that led to them more explicit, general, and perhaps most importantly, *generative*. That is, the insights and research foundations that led to these innovative techniques are put into a form (i.e., *design patterns*, to be discussed at length later) that helps assessment designers apply them more easily in a broader range of assessment situations and contexts.

3.4 Difficult Challenges in Assessing Knowledge Representations

The topic of extracting one's internal representation of external reality via knowledge representations and mental models is an extremely controversial one. As Sasse (1991) pointed out, this is because the 'internal' mechanisms about the way that the world is represented in our brain (and knowledge about these internal representations) are not directly observable. We know neither the structure of the real world nor the structure of the mental model that is supposed to represent it— we can only infer. The cognitive system has no access to reality except through perceptions, which are already internal. A question that often sparks debate among cognitive scientists concerns our capacity to objectively determine the nature, mechanisms, and accuracy of our internal mapping of the external reality. There is no inherent procedure to determine the accuracy of our mapping. Consequently, once knowledge models are extracted by one of the many available knowledge elicitation or extraction methods, it then becomes difficult to convince the scientific community that the externalized model is indeed a representation of the learner's mind. Researchers also argue that there are incongruencies and inconsistencies between mental models extracted using different techniques. For example, a study led by Evans et al., (2001) shows how different knowledge elicitation techniques produce different results—and, thus, dissimilar mental models. This, of course, adds fuel to the debate about the validity and reliability of knowledge elicitation techniques.

A case in point is the cognitive analysis Steinberg and Gitomer (1996) carried out in the course of designing Hydrive, an intelligent tutoring system (ITS) for aircraft hydraulics. A number of approaches could have been used to elicit expert mechanics' knowledge structures. Asking them to build a knowledge network would have produced a map of components and their relationships. Asking them to sort cards with components would have given insights into their understandings of dimensions of similarity, possibly in terms of functions and roles. Sorting cards with system problems would have produced dimensions of similarity perhaps with respect to causes and strategies. But Steinberg and Gitomer used the structured interview technique (specifically, task analysis) called PARI (Precursor, Action, Response, Interpretation; see Hall, Gott, & Pokorny, 1995) in order to elicit knowledge representations in terms of production rules for problem-solving, which they synthesized in the language of Newell and Simon's (1972) terminology of problem spaces, active paths, space-splitting, and so on. We would argue that *none* of the alternative elicitation methods would produce *the* correct internal knowledge representation of one of the experts. Each of them, however, produced an inscription — i.e., an external representation of behaviors produced from their internal representations. Regularities in their behaviors, hence the inscriptions, served admirably to structure the instructional content of the Hydrive ITS.

The job of the educational assessor, then, is not the same as the job of the cognitive researcher. The goal is not to obtain "the true" internal representation of a student's knowledge in some defined area, but rather to assess critical components of it. That is, does the student behavior suggest that critical pieces of information—such as concepts, relationships, or representational forms—are present or missing? Does he or she carry out procedures effectively? Does his or her behavior suggest gaps, inefficiencies, or misconceptions that might be rectified with certain kinds of instruction or practice? These educational questions can be framed and addressed using techniques such as those described earlier, without claims for the uniqueness or the fidelity at the level of implementation—as long as they lead to effective educational actions, as was the case with Hydrive project. The likelihood of this happening, of course, increases to the degree that the techniques for eliciting and approximating aspects of students' knowledge are consistent with both cognitive research on knowledge representation and the purposes and constraints of the assessment context. That is, one designs assessment tasks by choosing amongst techniques that provide somewhat different results, accepting that each offers an approximation that highlights certain aspects of internal representation and performance based thereupon; each too poses its own profile of costs, burdens on subject and analysts, reliability, and validity with respect to the purpose of the assessment. To ensure a more thorough and accurate reflection of underlying knowledge, it is recommended that multiple elicitation techniques be chosen and used at once.

4.0 Design Patterns

As an introduction to the remaining sections of this paper, we will briefly define the concept of a *design pattern*, with the intention of highlighting its relation to external and internal knowledge representations. *Design patterns* are schemas or structures for conceptualizing the components of assessment arguments and their interrelationships. *Design patterns* provide a structure that helps task designers access targeted aspects of knowledge and proficiency, including aspects of knowledge representation discussed in this paper. Contrary to the main topic of this paper thus far—i.e., internal knowledge representations—*design patterns* can be regarded as external knowledge representation tools that lay out explicitly the chain of reasoning for assessment tasks, from the inferences one wants to make, to the observations needed to ground them, and eventually to the situations or contexts that will elicit those observations. In the remainder of the paper, we will attempt to shed more light on the benefits and uses of *design patterns*; we then present and discuss five *design patterns* created specifically for assessing aspects of an individual's internal representations of knowledge.

4.1 Assessment Arguments

Relating what students do in particular situations to what they know or can potentially do more broadly is a defining feature of educational assessments. While much about the practice of educational assessment has changed during the past century—in the form of extensions, elaborations, refinement, and explication—foundational aspects of assessment arguments reveal a deeper kind of stability. Separating the *structure* of assessment arguments from their *substance* brings to light continuity in the forms of data, their interpretation, and inferences that can be drawn from them.

A famous quote by Messick (1994) helps to elucidate some of the essential components of an assessment argument:

> *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)*

The most important feature of this quote is the chain of reasoning Messick alludes to with regards to the process of assessment design. The rationale for each link grounds the validity of a claim about a student's knowledge, skills, or abilities (KSAs). Elements of Toulmin's (1958) argument structures can be used to relate these assessment principles to more concrete aspects of assessment design and development. The attributes of *design patterns* can be described in terms of their relationship to this elaborated argument structure.

Toulmin (1958) describes an *argumen*t as a link between *data* (i.e., what the learner generates in response to a task) and a *targeted claim* (i.e., the proposition we want to support with data—for instance, the level of student knowledge, skill, or ability) supported by a *warrant* (i.e., an abstract assertion or rationale that grounds the relation between the data and the claim), strengthened by *backing* (i.e., additional information that validates the warrant, including theory, experience, empirical studies, prior research), and tempered by a *rebuttal* or *rebuttals* (i.e., an assertion that limits or situates or contextualizes the scope of the claim).

For example, in an achievement assessment scenario, the targeted claim might concern the degree to which students understand basic elements of kinematics in the domain of physics. Students with deep conceptual understanding about basic elements of kinematics tend to perform well on tasks that require them to recognize and/or produce alternative expressions of the same underlying meaning given by a target (e.g., recognizing preservation of meaning when principles of velocity are represented using a variety of modalities including formula, a velocity vs. time graph, or a diagram). This is the warrant. The backing for the warrant is the empirical support for the importance of: a) multiple representations to learning, and b) intricate, hierarchical, organized connectivity between related concepts and sub-concepts, relations and interrelations. Learners' responses (i.e., the data) to meaning equivalent conceptual problems are used to infer the degree of understanding about the content being assessed. A possible rebuttal to this argument could be that a learner may have provided the correct answer to the problem as a result of a lucky guess. This rebuttal may, in turn, be considered highly improbable since students also were asked to explain the reasoning behind their answer. Nevertheless, it is probable that other valid rebuttals could be projected; any conclusions about our claims would have to be qualified in light of them. As Schum (1994) indicated, it is always necessary to establish the credentials of evidence, its relevance, credibility, and force.

4.2 The Structure of Design Patterns

Design patterns are one component of the Principled Assessment Designs for Inquiry (PADI) project. PADI is a conceptual framework used for providing well thought out navigation in the design of inquiry tasks, coordination of and provision of tools to facilitate them. Although the PADI system focuses on the entire range of the process from purpose of the assessment to the substance of the arguments (including the *design patterns*) to the technical details of operational elements and delivery systems, only the role, structure, benefits, and application of the *design pattern* concept will be further explored here.

Design patterns fulfill a number of very important roles in the assessment design process. First and foremost, *design patterns* integrate knowledge about what is important to assess within the structure of an assessment argument, in a format that readily guides task creation and assessment implementation. Each *design pattern* sketches what amounts to a narrative structure concerning, but not limited to, the knowledge, skill, or abilities (KSAs) one wants to address (i.e., Focal KSAs), other knowledge, skills, and/or abilities that may be required by the task (i.e., Additional KSAs), the kinds of observations that can provide evidence about the acquisition of the specified KSAs (i.e., Potential Observations), features of task situations that can evoke the desired evidence about the target KSAs (i.e., Characteristic Features), and features can be varied in order to shift the difficulty or focus of tasks (i.e., Variable Features). See Table 3 for more details. *Design patterns* allow communication between educators and assessment designers, two important stakeholders in the process of assessment design, in a non-technical but meaningful way that facilitates the creation of assessment tasks (Messick, 1994).

Table 3. Attributes of a PADI Assessment Design Pattern

ATTRIBUTE	DEFINITION
Title	A short name for referring to the design pattern
Summary	Overview of the kinds of assessment situations students encounter in this design pattern and what one wants to know about their knowledge, skill, and abilities
Rationale	Explanation of why this is an important aspect of scientific inquiry; the underlying warrant that justifies the connection between the targeted inferences and the kinds of tasks and evidence that support them
Focal Knowledge, Skills, & Abilities	The focal knowledge, skills, and/or abilities targeted by this design pattern
Additional Knowledge, Skills, & Abilities	Other knowledge, skills, and/or abilities that may be required by tasks motivated by this design pattern
Potential Observation(s)	Some possible things that one could see students doing that would provide evidence about their focal knowledge, skills, and/or abilities
Potential Work Product(s)	Student responses or performances (e.g., written product or spoken answer) that can hold clues or provide evidence about the focal knowledge, skills, and/or abilities
Potential Rubric(s)	Rules and techniques used or adapted for evaluating or scoring work products
Characteristic Feature(s)	Features of an assessment task or situation that can evoke the desired evidence about the focal knowledge, skills, and/or abilities
Variable Feature(s)	Features of an assessment task or situation that can be varied in order to shift the difficulty or focus of tasks
Exemplar Tasks	Sample assessment tasks that are instances of this design pattern
References	Research and literature that provide backing for this design pattern

There are three key benefits of using *design patterns*. Note, that since these are discussed in detail elsewhere, we will provide only a brief overview (see DeBarger & Riconscente, 2005, PADI Report 8). The first benefit of *design patterns* is that they *facilitate decision-making about assessment design*. There are many macro- and micro-level decisions to be made in assessment design. Instead of making the decisions after or during assessment design, the pre-thinking and decision-making that goes into the creation of *design patterns* are intended to identify for the assessment designer which decisions have already been made and which still need to be made.

Design patterns make explicit the *thought process* that comes into play in designing assessments, explicitly identifying the goals of the assessment, how one will measure them, what Potential Work Products look like, etc., so that details of the driving force behind the assessment are clear, minimizing compromised inferences about student understanding or ability, or what the assessment results mean. The second key benefit of using *design patterns* is that it helps to *explicate the assessment argument*. In laying out the design decisions, a crucial consideration is the assessment argument or the line of reasoning that will ultimately connect the assessment task to the inferences we wish to make about students' knowledge, skills, or abilities. *Design patterns* explicate, articulate, and make explicit the details of the assessment argument by the decisions expressed in each attribute of the *design pattern*. A third benefit of using *design patterns* is its *flexibility; design patterns* can address a range of psychological perspectives on learning that provide the rationale behind the development of different assessments (while *design patterns* guide important decisions in the design process, the content of these decisions remains open to the perspective and purpose of the assessment). They can also vary in generality and scale (DeBarger & Ronscente, 2005). For more detailed information about the nature of *design patterns*, refer to the PADI Technical Report 1 (Mislevy et al., 2003).

5.0 Five Design Patterns of Assessing Knowledge Representations

Design patterns can be created in a number of ways. One way is to start from an existing assessment and work backwards to extract a more general *design patterns* that may be used to generate similar kinds of assessment. Another strategy consists of beginning with a set of learning outcomes to be included in the Student Model and proceeding from there to identify appropriate Potential Observations, Work Products, Rubrics etc. In this paper we have created *design patterns* using both strategies. Out of the five *design patterns* below, one was created using the first strategy (i.e., working backwards): Recognizing and Producing Equivalence of Meaning. The other four *design patterns* were created using the second strategy (i.e., working forwards): Building Knowledge Representations; Updating & Correcting Knowledge Representations; Using Knowledge Representations; and Re-expressing Data. These *design patterns* take on a cognitive perspective in trying to assess aspects of knowledge representation detailed in earlier parts of this paper.

In the sections that follow, we will present and discuss five *design patterns*, addressing their attributes and providing examples that constitute an instantiation of the ideas within the *design pattern*, as determined by the constraints and purposes of their intended use.

5.1 Design Pattern for "Building Knowledge Representations"

Table 4 summarizes a *design pattern* to guide the construction of tasks that aim to assess aspects of how the learner's knowledge representations are built, including how concepts are organized and related. The attributes of this and all *design patterns* are organized according to the elements of an assessment argument. They suggest design choices for a task author to consider, motivated jointly by research on knowledge representations and considerations of assessment design.

Summary & Rationale: The process of comparison and organization of concepts and sub-concepts by their structure, function, or relation is an important aspect of mental model building and understanding. The "Building Knowledge Representations" *design pattern* has dual function. It serves to elicit details about learners' current understanding or perception of the hierarchical functional/structural relationships and interrelations between group elements; in addition, it also functions to help further elucidate that initial understanding by requiring that learners explicitly think about the relations between facts or concepts as they work through the tasks.

In this *design pattern*, learners are given a list of concepts about a domain or subset of a domain and are asked to divide, sort, compare, and organize them into groups and subgroups according to their understanding of the structures, relations, and interrelations amongst them. Once the division and grouping is completed, learners are required to "name" their groups and to explain the underlying common structure or relationship between group elements. This *design pattern* incorporates the use of a number of knowledge elicitation techniques (described earlier) including interview and conceptual techniques, and in particular, knowledge proximity measures primarily concerned with assessing the degree of relatedness between two or more domain concepts, principles, or ideas.

It is important to point out that *design patterns* can be coarse or fine–grained — that is, either broad (i.e., providing an umbrella for many other specific ones) or very specific. The "Building Knowledge Representation" *design pattern* is an example of a specific one, suggesting that there can be other *design patterns* with a similar goal of designing tasks for assessing the building of knowledge representations.

Table 4. Design Pattern for Building Knowledge Representations

ATTRIBUTE	**VALUE (S)**	**COMMENT (S)**
Title	Building Knowledge Representations	
Summary	In this design pattern, learners are given a list of concepts about a domain or subset of a domain and asked to divide, sort, compare, and organize them into groups and subgroups according to their understanding of the structures, relations, and interrelations amongst them. Once the division and grouping is completed, learners will be required to 'name' their groups, explaining the underlying common structure or	Concepts may include facts, principles, or ideas.

	relationship between group elements. The quality of learners' response will reveal details about, and help to further build, an understanding of the hierarchical functional (and/or structural) relationships and interrelationships between the elements of a system.	
Rationale	Knowledge representations are characterized by their ontologies (i.e., what is left out being as important as what is included); the processes they support; and the fundamental concepts that ground them. The process of comparison and organization of concepts and sub-concepts by their function or relation is an important aspect of mental model building and understanding.	
Focal Knowledge, Skills, & Abilities	Ability to identify meaningful patterns of relationship (i.e., with regards to structure or function) between given concepts, sub-concepts, or sets thereof. Ability to *categorize* and *organize* concepts and sub-concepts with varying levels of relations, links, or interconnections. Ability *to pinpoint* and *articulate the underlying common structure* or relationship between group elements; *and their reasons/rationale* behind various groupings. Ability to identify when structures, relations, or interrelations between concepts or sets of concepts are related by *superficial similarities or deep principles*; when the similarities or differences are relevant or irrelevant. Level of *conceptual knowledge* and/or understanding about the intricate network of hierarchically connected, organized ideas, relations, and interrelations between various levels of domain concept and sub-concepts.	
Additional Knowledge, Skills, & Abilities	Ability and efficiency in performing cognitive tasks such as sorting, comparing, organizing, and/or grouping data or evidence. Underlying knowledge activation and retrieval; memory and information processing.	

	Relevant domain knowledge (e.g., familiarity with each concept on the list). Ability to manage, monitor, and reflect upon thinking during comparison, categorization, and organization process and to allocate mental resources efficiently.	
Potential Observation(s)	Pattern and efficiency of comparing, connecting, equating, and grouping various super-ordinate and sub-ordinate concepts according to function, structure, or relationships (accuracy or partial accuracy). Content, scope, and nature of categories (what elements are included; are they grouped by deep principles or surface characteristics, functional relationships, or structure etc.). Number of categories. Distinctness of categories. Accuracy of determination of what concepts/sub-concepts are important to include together. Accuracy of representations of relations among concepts/sub-concepts. Time it takes to complete the task. Quality (also relevancy, accuracy, correctness) of a) *categorizations;* b) *written/verbal defense* of categories; c) *graphical* network representations.	
Potential Work Product(s)	Student grouping/ categorization of concepts and sub-concepts. Student identification of meaning behind choice of categories and the grouping of certain elements (i.e., reason or rationale for placing different concepts within the same category). Student written/verbal explanations of the identified relationships and interrelations between the elements on the list. Student written and verbal explanations or defense for grouping/categories.	

	Student graphical 'network' representations of super-ordinate and sub-ordinate relationships and interrelationships between concepts (e.g., hierarchical tree diagram; Venn diagram and/or Flow charts) Notes taken by the learner during the process (e.g., about relationship or meaning between concepts).	
Potential Rubric(s)		Scoring would be based on the accuracy and completeness of the groupings; verbal defense; and graphical network, as compared with characteristic groupings of an expert response.
Characteristic Feature(s)	Tasks that ask student to categorize, organize, and/or represent various concepts and sub-concepts according to their structures, relations, or interrelations.	
Variable Feature(s)	Varying the *content, scope, complexity, and nature* of the concepts and sub-concepts (e.g., inclusion of concepts with conditional or context dependent relationships and/or inclusion of abstract concepts creates more difficulty). *Number* of concepts/sub-concepts given in list. *Working memory* and *cognitive processing load* required. *Level of abstraction* between concepts or sub-concepts. *Complexity of relations* between concepts or sub-concepts. Complexity of the *real world system* being modeled. Whether to use a *list of concepts* or *list of problems* to be categorized. Whether to use the task as a *learning tool* or formal *assessment tool.*	

	Whether the task should be an *individual or group* activity. Whether to maximize or limit the *level/complexity/ range (and integration)* of conceptual knowledge required to solve problem.	
Exemplar Tasks	Variants of Chi and colleagues categorization tasks in Physics (1981).	
References	Chi, M. T. H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5(2),* 121-152. Geiwitz, J., Kornell, J., & McCloskey, B. (1990). An Expert System for the Selection of Knowledge Acquisition Techniques. Technical Report 785-2. Santa Barbara, CA: Anacapa Sciences. Cordingley, E. S. (1989). Knowledge elicitation techniques for knowledge-based systems. In D. Diaper (ed.), Knowledge Elicitation: Principles, techniques, and applications. John Wiley & Sons. Fincher, S., & Tenenberg, J. (2005). Making sense of card sorting data. *Expert Systems, 22(3),* 89-93.	

The Focal Knowledge, Skills, or Abilities (i.e., KSAs) of this *design pattern* are particular aspects of mental model building and understanding, in particular, the ability to compare and organize domain concepts and sub-concepts by their structure, function, relations at varying levels of relatedness, and to be able to explicitly pinpoint and articulate the underlying common structure or relationship (whether deep or surface) and their reason/rationale for the various groupings. Depending on the purpose of the assessment, a designer may choose to focus more on some KSAs than on others, or they may wish to address them separately and specifically or as an ensemble.

Additional KSAs are skills, other than the target, that may be required in a task developed from this *design pattern*. This attribute of *design patterns* call the assessor's attention to design choices that may elicit or minimize the learner's ability to complete the task. An assessment designer must consider, for instance, whether all task content should be familiar to learners so that it is not a confounding factor or whether familiarity is unnecessary. For instance, is it necessary for the learner to be familiar with each concept on a given list? Should a brief definition of all concepts also be given? Familiarity with the task type (i.e., having to sort, compare, organize, group data or evidence) and metacognitive needs to manage, monitor, and reflect upon thinking during the task and to allocate mental resources efficiently are Additional KSAs for this *design pattern*.

Variable Task Features are characteristics of the task that the assessment designer can choose to vary in order to shift the difficulty or focus of tasks. For the "Building Knowledge Representations" *design pattern*, one can choose to vary, among other things, the *content, scope, complexity, and nature* of concepts, sub-concepts, and their interrelations (e.g., inclusion of conditional or context dependent relationships; the level of abstraction for instance, first- and second-order concepts); and the number of concepts and/or sub-concepts given in the list. For instance, in one of the simpler versions of a task from this *design pattern*, the learner could be asked to sort a series of six cards, denoting only first-order concepts (e.g., cat, dog, flower, tree, shark, goldfish), according to similarity (i.e., resulting in three groups: animals; plant; fish). In a more complex version of a task from this *design pattern*, the learner could be asked to sort *more* (e.g., 10) cards, denoting both *first- and second-order* concepts (e.g., cat, dog, Siamese cat, flower, tree, shark, goldfish, golden retriever, conifer, perennial) into groups and sub-groups and to correspondingly label the groups, sub-groups, and interrelations accordingly. The learner's ability to identify subgroups within the major group (e.g., Siamese cats as a type of cat, and cat as a type of animal) reflects a greater degree of differentiation amongst concepts that could be targeted by the more complex version of the sorting task.

Work products that can be called for in tasks from the “Building Knowledge Representation” *design pattern* include information about the learner’s organization of knowledge and their idiosyncratic process or justification for the sorting. Information about the learner’s knowledge organization can take various forms, which encompass as examples the resulting sort or categorization of concepts and sub-concepts. Work Products that focus on the process can include the learner’s explanations (and defense) of the identified relationships and interrelationships between elements of the list, which may be written or verbal (including conversations students have with themselves as they sort); notes taken during the construction process about the relationship or meaning between concepts, and other representations of the learner’s graphical network of representations such as hierarchical tree diagrams, Venn diagrams, or flow charts.

Several kinds of Potential Observations can be evaluated from these Work Products. Regarding information about the learner’s knowledge organization, Potential Observations include the content, scope, complexity, and nature of categories (e.g., what elements are included, are there first- and second-order categories, are the elements grouped by deep or surface characteristics etc); the pattern and efficiency of comparing, connecting, equating, and grouping various super-ordinate and sub-ordinate concepts; and the number, distinctiveness, relevancy, accuracy, and quality of categorizations. Potential Observations regarding process include time efficiency; accuracy and completeness of the learner’s explanations/defense of categories and interrelations; accuracy and completeness of graphical representations; and metacognitive awareness of the process.

Potential Rubrics developed for the “Building Knowledge Representations” assessments provide the means by which Work Products are evaluated and translated into observations germane to the Focal KSAs. Scoring for assessment tasks under this *design pattern* would be based on the accuracy and completeness of the categories; verbal defense; and graphical network/student’s explanations of relations, as compared with the characteristic groupings of an expert response.

An exemplar task instantiating this *design pattern* and used to illustrate it, is Chi and colleagues’ (1981) card-sorting task. Chi and colleagues (1981) conducted several studies which serve to illustrate the many attributes of this *design pattern*. In Chi’s task, for example, physics problems written on index cards were given to subjects comprised of both students and experts. In the first study, subjects were asked to group the physics problems according to similarity of solution (i.e., Work Product). They examined the nature of the sorts produced, including the dimensions used to define categories, the assignment to categories, and the order in which the various dimensions are produced (i.e., Potential Observations). Chi and colleagues found that the sorts produced reflected a deep versus superficial distinction. Specifically, students (or novices) categorized problems in terms of surface characteristics such as the physical terminology or physical configuration mentioned in the problem. For instance, problems involving inclined planes were grouped together regardless of whether or not they could be solved by applying the same principle. Experts, on the other hand, categorized physics problems by the primary principles, ideas, concepts, or procedures used in the solution to the problem. In a related study (Chi et al., 1981), both novices and experts were asked in an open-ended interview format to tell all they could about problems classified in each grouping (i.e., Work Product). The interviews were analyzed for mention of general principles or surface details (i.e., Potential Observation). Again, Chi and colleagues found that experts used problem solving schemata which circulated around general principles and contained an ordered hierarchy of conditions while novices had a lot to say about the superficial details of the categories. In a final study, Chi and colleagues (1981) modified the features of the task yet again, this time asking experts and novices to think aloud — that is, articulate how they would solve the physics problems while they were actually solving it (i.e., Work Product). Potential Observations for the articulation include clarity, accuracy, and whether it centered on deep underlying principles or surface features of the task. Again, they found evidence suggesting that experts first determined the underlying principles while novices jumped into solving the problem using equations.

5.2 Design Pattern for "Updating, Elaborating, & Correcting Knowledge Representations"

Table 5 summarizes a *design pattern* to guide the construction of tasks that aim to assess aspects of how the learner updates or elaborates incomplete knowledge representations or corrects faulty ones.

Summary & Rationale: In this *design pattern*, learners are presented with data, a situation, or real life simulations only *partially consistent* with existing, already constructed, knowledge representation (s). This partially consistent information is inadequate for problem solving. Learners are asked to dynamically interact with the data, use problem solving, and reasoning skills (first based on their existing and then updated/corrected knowledge representations) to answer probing questions, test hypotheses, and predict future states. In the midst of completing tasks, learners evaluate the correspondence between their mental models and the real world data. In some cases, learners' current mental model may not be as "accurate" or "complete" as it should be and consequently may be insufficient for solving problems at hand. Activities motivated by this *design pattern* allow learners to recognize and evaluate discrepancies between their current knowledge and the external world via incorrect predictions about the behavior of the system (due to incorrect or incomplete rules etc). Tasks generated from this *design pattern* also will provide the data needed for making necessary extensions, corrections, and/or adjustments to faulty/deficient mental models.

Note that in this *design pattern*, current knowledge is inadequate for problem solving; this *design pattern* is contrasted with the "Using Knowledge Representations" *design pattern* (to be described later), which requires learners to apply existing mental models to solve problems for which their current knowledge is adequate.

Table 5. Design Pattern for Updating, Elaborating, & Correcting Knowledge Representations

ATTRIBUTE	VALUE (S)	COMMENT (S)
Title	Updating, Elaborating, and Correcting Knowledge Representations	
Summary	In this design pattern, learners are presented with data, a situation, or real life simulations only *partially consistent* with existing, already constructed, knowledge representation (s). Learners are asked to dynamically interact with the data, use problem solving and reasoning skills (based on their existing and then updated/corrected knowledge representations) to answer probing questions, test hypotheses, and predict future states. Dynamic interaction with the data allows the learner to evaluate the correspondence between their mental model and the real world data. Their identification of anomalies results in extensions or revisions to their current mental model.	This design pattern is geared at facilitating the elaboration, correction, and/or updating of existing knowledge representations once a discrepancy or anomaly between the extracted information from the external world and incorrect paths of the internal mental model are detected (e.g., misconceptions). In this design pattern, current *knowledge is inadequate* for problem solving. This design pattern is contrasted to the "Using Knowledge Representations" design pattern, which requires students to apply existing

		mental models to solve problems for which that *knowledge is adequate*.
Rationale	In some cases, a learner's current mental model may not be as 'accurate' or as 'complete' as it could/should be; and may be insufficient to solve problems at hand. Activities motivated by this design pattern allow learners to recognize discrepancies between their current knowledge and the external world via incorrect predictions about the behavior of the system (due to incorrect or incomplete rules) etc. Tasks generated from this design pattern will provide data needed for making necessary corrections, elaborations, and adjustments to faulty/deficient mental models.	
Focal Knowledge, Skills, & Abilities	Ability to use a given dataset, situation, or simulation to reason through the concepts and relationships of represented knowledge, in order to solve problems, answer questions, test hypotheses, or make predictions. Ability to identify salient features of given data. Ability to recognize the relevance of mental model/knowledge to particular data, situation, or simulation scenario. Ability to identify discrepancies between reasoning ensuing from existing knowledge representation (due to misconceptions or incomplete mental models) and evidence provided by real world data. Ability to use new/discrepant information to update, elaborate, and/or correct existing knowledge representations (i.e., account for anomalous data).	
Additional Knowledge, Skills, & Abilities	Familiarity with and understanding of the process of scientific inquiry (e.g., procedure for testing hypothesis; reasoning and making inferences). Familiarity with various task type (situations, data, or simulations). Underlying knowledge activation and retrieval; memory and information processing.	

	Relevant domain knowledge. Ability to manage, monitor, and reflect upon thinking and to allocate resources efficiently. *	*Note that this could be developed into a design pattern of its own
Potential Observation(s)	Use knowledge to reason, solve problems, test hypotheses, and make predictions (relevancy, accuracy, and quality of response). Accuracy and completeness of mapping between elements and/or relationships of represented knowledge and real world situation, data, or simulated task. Identification of ways that knowledge does not match the data/or situation. Making appropriate revisions to mental model after observing discrepancies. Quality of the basis on which students decide that a revised model is correct and/or complete. Quality of student's modifications to current mental model. Dynamic and thoughtful interaction with task.	
Potential Work Product(s)	Correspondence mapping between elements and/or relationships of represented knowledge and real world situation, data, or simulated task. Correspondence mapping between elements and/or relationships of adjusted mental model and real world situation, data, or simulated task. Learners' solutions to given problems, based on interaction with task (constructed or selected response). Learners' response to tested hypotheses (constructed or selected response). Learners' predictions/inferences (constructed or selected response).	

	Learners' written and verbal 'think aloud' explanations of reasoning, answers to given problems, hypotheses, predictions (based on initial flawed model and revised model). Learners' written or verbal explanation or reasoning for elaborating, correcting or otherwise revising the current mental model. Tracing or monitoring of online activity with simulation model.	
Potential Rubric(s)		
Characteristic Feature(s)	Real-world situation, data or simulation and represented knowledge, for which details of correspondence need to be fleshed out. Focus is on using hierarchically-based knowledge about relations and interrelations between concepts and sub-concepts to interact with, problem solve, and reason about the task. Addresses correspondence between, and application of, knowledge representations to the situation or task.	Pattern in the data **contradicts** the learners' model.
Variable Feature(s)	*Varying task* to be used for fleshing out details of correspondence (e.g., real word situation, data or simulation). Familiarity of context of data. *Level and nature* of represented knowledge required to solve problem. Degree of *data contradiction* with the existing model (core?). *Plane of contradiction* (at level of components, connections, relations, functions, structures). *Working memory* and *cognitive processing load* required. Whether to maximize or limit the *level/complexity/ range (and integration)* of conceptual knowledge required to solve problem. Level of prompting/scaffolding required Individual work, with a partner, or as a member of a group.	

Exemplar Tasks	Interaction with, and interpretation of dynamic, online simulated graphs such as a velocity versus time graph. Simulation allows the manipulation of parameters and real time observation of the consequences of this manipulation. Web-based Inquiry Science Environment (WISE) website for examples. Concord Consortium's simulation examples.	
References	Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. Journal of Experimental Psychology: Learning, memory, & cognition, 19(5), pp. 1061-1070. diSessa, A. A. (1993). Toward an 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. Sigel, I. E. (1999). Approaches to representation as a psychological construct: A treatise in diversity. In I. E. Sigel (Ed.), Development of mental representation: Theories and applications (pp. 3-12). Mahwah, NJ: Erlbaum.	

The Focal Knowledge, Skills, or Abilities (i.e., KSAs) of this *design pattern* are particular aspects of updating, elaborating, and correcting existing mental models — in particular, the ability to use a given data set, situation, or simulation to reason through the concepts and relationships of represented knowledge in order to solve problems, answer questions, test hypotheses, or make predictions; the ability to identify salient features of the given data; the ability to identify discrepancies between reasoning ensuing from existing knowledge representation (due to misconceptions or incomplete mental models) and evidence provided by real world data; and the ability to use discrepant information to update, extend, and/or correct existing mental models. These Focal KSAs are all dependent on the learner's ability to recognize the relevance of a mental model/knowledge representation to particular data, situation, or simulation scenario.

Additional KSAs that may be required in tasks developed from this *design pattern* include, for instance, the learners' familiarity with and understanding of the process of scientific inquiry — that is, the extent of their knowledge about the procedure for forming or testing hypotheses, using data to refine knowledge, or for that matter, the procedure for reasoning and making inferences from data. Also important is their familiarity with using real life external situations to reason. Important questions in this regard include: Do learners have the relevant domain knowledge, and if they do, are they able to activate and/or retrieve this information from where it is stored at the precise time that it is needed? Can they manage, monitor, and reflect upon thinking and allocate resources efficiently? Additional KSAs are important factors to consider when designing and interpreting results from assessment tasks. An analysis of Additional KSAs helps to ensure that it is the target or focal knowledge that is influencing student performance on tasks and not something external, such as unfamiliar task type, information processing deficiencies, or inadequate domain knowledge.

Variable Task Features: For the "Updating, Elaborating, and Correcting Knowledge Representations" *design pattern*, one can choose to vary, amongst other things: the task used for fleshing out details of the correspondence (specifically, whether it is a real world situation, compiled data, or computer simulation); whether the context of data presentation is familiar or unfamiliar (i.e., has the student used computer simulations before); the level and nature of represented knowledge required to solve the problem (e.g., involving deep levels of the hierarchy or perhaps more than one mental models); the degree of data contradiction with the existing model (e.g., a discrepancy central to the mental model would be easy to identify relative to a discrepancy of a conditional feature which would be more difficult, requiring deeper knowledge of the particular aspect of the domain); the plane of contradiction (i.e., whether the discrepancy is at the level of components, connections, relations, functions, or structures); the degree of prompting or scaffolding required to complete the task; and whether the task should be administered to an individual, a small group, or a large group.

Work Products that can be called for in this *design pattern* include: information about learners' response to the task; information pertaining to learners' recognition of the mismatch (and later match) between their current (and/or adjusted) mental models and the real world situation, data, or simulated task; and non-verbal information. Work Products related to *learners' response to the task* includes solutions to given problems, responses to tested hypotheses, predictions, inferences, and reasoning, based first on their flawed model and then their revised model. Work Products related to *learners' mismatch* (and later match) between the external world versus internal model can take various forms, which encompass as examples, learners' correspondence mapping between elements and/or relationships of represented knowledge and real world situation, data, or simulated task, first for the current mental model and then for the adjusted mental model; and learners' explanations for reasoning behind the mismatch and their reasoning for elaborating, correcting, or revising the existing mental model. Note that the hypotheses, request for predictions, and so forth, may be in response to a selected response assessment item (e.g., multiple choice), open-ended, or even self prompted, depending on the degree of scaffolding that is included in the task (i.e., Variable Feature). Other *non-verbal* *process* tracing information can be gathered from online activity, as in the case with a computer-based simulation model, including number of key strokes, incorrect responses, time lags, and eye movements.

Several kinds of Potential Observations can be evaluated from these Work Products. Regarding *learners' response to the task*, Potential Observations include the relevancy, accuracy, and quality of their response when asked to use their knowledge to solve problems, make predictions, hypotheses, and inferences. Regarding *learners' mismatch*, Potential Observations include the quality of the basis (and ways) in which learners' decide that their current model does not match the data/or situation; the accuracy and completeness of their mapping between elements in the external data and their internal mental model; the reasoning and quality of revisions made to their mental model after observing discrepancies; and the quality of the basis on which learners' decide that a revised model is now correct and/or adequate.

Examples of tasks that exemplify this *design pattern* can involve the use of simulations. Simulations present learners with rich task environments that are geared at modeling systems in the natural world; Because simulations are dynamic and can be manipulated (with real time consequences), learners can demonstrate their abilities to engage in active inquiry. The first example from the Physics Education Technology Project is called "The Moving Man." Here the task is to explore and learn about velocity, position, and the association between the two in relation to time. After forming a set of hypotheses and predictions based on their current knowledge, learners' use this simulation applet to manipulate and test their hypotheses and predictions using real data. Some of the variables that can be manipulated in the simulation include speed and direction of movement of an object (i.e., back or forward). Learners then discuss the principles needed to explain the data, consequently elaborating, revising, and/or correcting their current mental models if deemed necessary. See a screen shot of "The Moving Man" task in Figure 2.

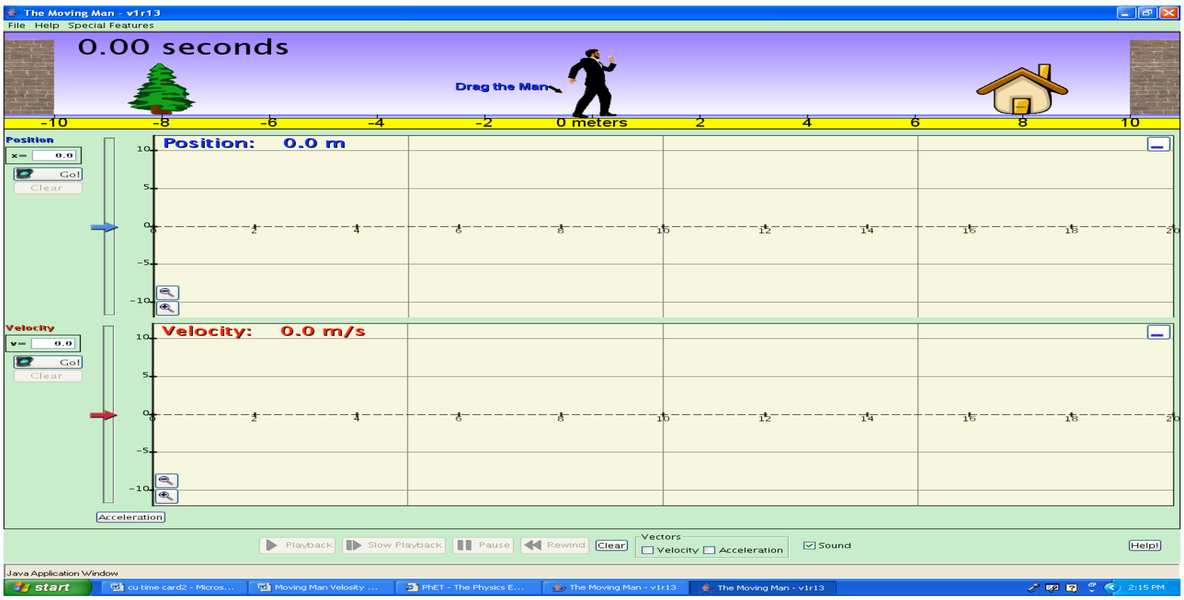

Figure 2. The Moving Man Task

Another task example that instantiates this *design pattern* originates from the Web-based Inquiry Science Environment (WISE) project — in particular, the "Probing your Surroundings" task. This task is designed to assess learners' knowledge about thermal equilibrium in the context of the temperature of objects around them. Similar to "The Moving Man" task, after making predictions and answering probing questions, learners are given the chance to experiment and collect data. They are prompted to notice patterns in the data gathered and to develop principles. They are also prompted to evaluate whether and why the data pattern matched (or did not match) their predictions. Finally, they are asked to discuss the data pattern with others in the class, which includes defending their own principles, giving evidence or examples to support it, and commenting on other learner's principles.

In another example, dealing with particle speeds and temperature, learners are asked about the concept of temperature — specifically, how the distribution of particle speed is related to temperature. Learners then use the simulation to test their ideas. In the simulation applet, the thermometer records the temperature of the gas with the mean particle speed (MPS) and pressure (Pres) given. There is also a graph window to the right that shows the change in the velocity of the particle in the balloon as the temperature increases. In this particular version of the simulation, the learner can manipulate only the temperature of the gas inside the balloon. This example comes from the University of Oregon's Physics Department. See a screen shot of this simulation in Figure 3.

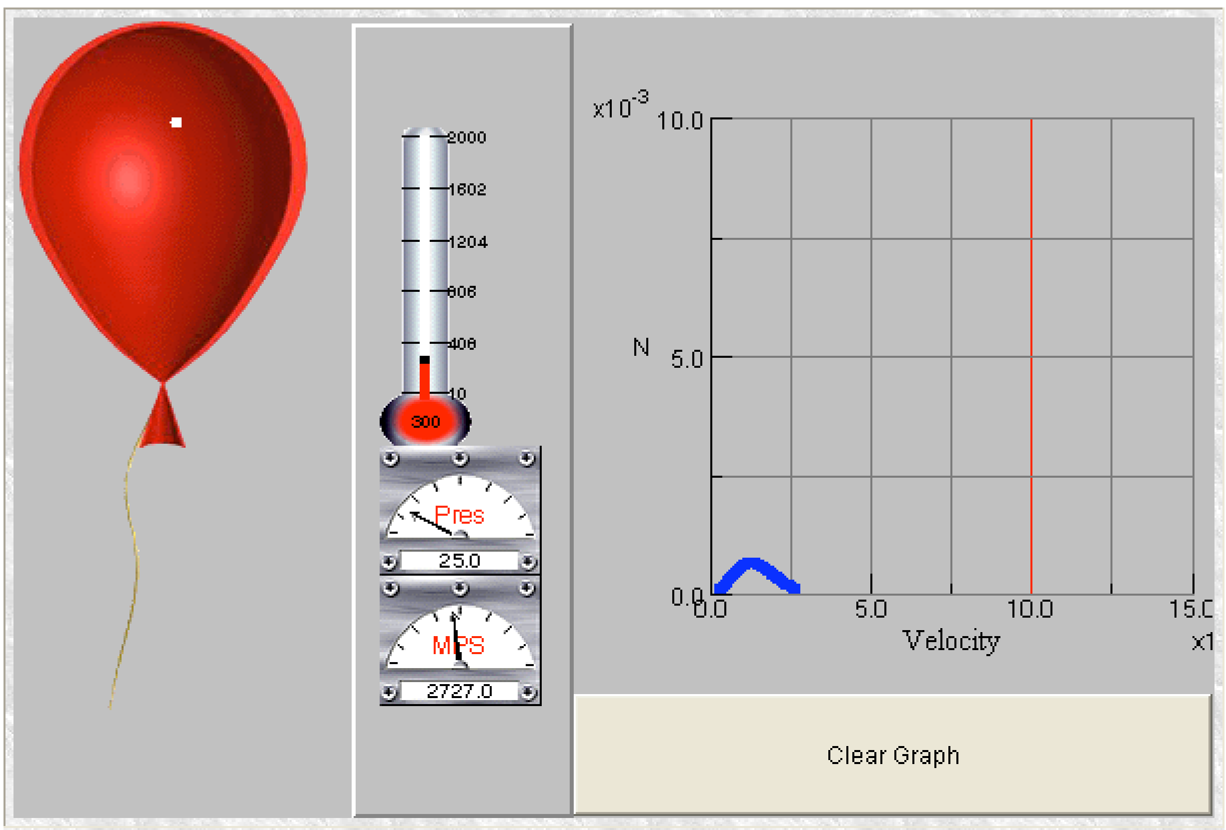

Figure 3.   Particle Speed and Temperature Task

5.3 Design Pattern for "Using Knowledge Representations"

Table 6 summarizes a *design pattern* to guide the construction of tasks that aim to assess the usage of knowledge representations.

Summary & Rationale: Activities motivated by this *design pattern* allow learners to extract key functions and relationships from their mental model; apply this knowledge to test hypotheses, solve problems, and make predictions and inferences. This process allows for the strengthening of existing mental models. In this *design pattern*, learners are provided with data, a situation, or real life simulations *consistent* with existing, already constructed, knowledge representation (s). Learners must dynamically interact with the data, use problem solving and reasoning skills (based on their existing knowledge representations) to answer probing questions, simulate the system mentally, test hypotheses, and predict future states. Correspondence between the real world data and existing mental models is sufficient for solving the problems.

Note how this design pattern is parallel to the one before it, "Updating, Elaborating, and Correcting Knowledge Representations." The main difference between the two is that here the learner's existing mental model is adequate for completing the task — that is, his or her mental model does not need to be revised, corrected, or elaborated.

Table 6. Design Pattern for Using Knowledge Representations

ATTRIBUTE	VALUE (S)	COMMENT (S)
Title	Using Knowledge Representations	
Summary	In this design pattern, learners are provided with data, a situation, or real life simulations *consistent* with existing, already constructed, knowledge representation (s). Learners must dynamically interact with the data, use problem solving and reasoning skills (based on their existing knowledge representations) to answer given questions, simulate the system mentally, test hypotheses, and predict	This design pattern is geared at how people make use of internal knowledge representations once they are constructed. One type of problem solving that involves learning is the application of already

	future states.	constructed knowledge to data, which are *consistent* with that knowledge.
Rationale	Activities motivated by this design pattern allow learners to extract key functions and relationships from their mental model; apply this knowledge to test hypotheses, solve problems, and make predictions and inferences. This process will allow for the strengthening and elaborations of existing mental model.	Elaborations in terms of making the mental model more concrete and perhaps detailed.
Focal Knowledge, Skills, & Abilities	Ability to use a given dataset, situation, or simulation to reason through the concepts and relationships of represented knowledge, in order to solve problems or make predictions. Ability to identify salient features of given data. Ability to recognize relevance of knowledge to particular data, situation, or simulation scenario (knowing when and how to apply that knowledge).	
Additional Knowledge, Skills, & Abilities	Familiarity with and understanding of the process of scientific inquiry (e.g., procedure for testing hypothesis; making inferences). Familiarity with various task type (situations, data, or simulations). Underlying knowledge activation and retrieval; memory and information processing. Relevant domain knowledge. Ability to manage, monitor, and reflect upon thinking and to allocate resources efficiently. *	*Note that this could be developed into a design pattern of its own
Potential Observation(s)	Use given data and represented knowledge to solve problems, test hypotheses, make and explain predictions (relevancy and accuracy). Accuracy and completeness of mapping between elements and/or relationships of represented knowledge and real world situation, data, or simulated task. Dynamic and thoughtful interaction with task.	

Potential Work Product(s)	Correspondence mapping between elements and/or relationships of mental model (s) and real world situation, data, or simulated task. Correspondence mapping between elements or relationships of overlapping mental models (if more than one is involved). Learners' solutions to given problems or questions, based on interaction with task (constructed or selected response). Learners' response to tested hypotheses (constructed or selected response). Learners' predictions/inferences (constructed or selected response). Learners' written and verbal 'think aloud' explanations of reasoning, answers to given problems, hypotheses, and predictions. Tracing or monitoring of online activity with simulation model.	
Potential Rubric(s)		
Characteristic Feature(s)	Real-world situation, data or simulation and represented knowledge, for which details of correspondence need to be fleshed out. Focus is on using hierarchically-based knowledge about relations and interrelations between concepts and sub-concepts to interact with, problem solve, and reason about the task. Addresses correspondence between, and application of, knowledge representations to the situation or task.	Patterns in the data **do not** contradict the mental model/knowledge represented.
Variable Feature(s)	*Varying task* to be used for fleshing out details of correspondence (e.g., real word situation, data or simulation). Familiarity of context of data. Using single or multiple mental model(s) to reason through data, situation, or simulation. *Working memory* and *cognitive processing load* required. Whether to maximize or limit the *level/complexity/ range (and integration)* of conceptual knowledge required to solve	

	problem. Level of prompting/scaffolding required Individual work, with a partner, or as a member of a group.	
Exemplar Tasks	Interaction with and interpretation of dynamic, online simulated graphs such as a velocity versus time graph. Simulation allows the manipulation of parameters and real time observation of the consequences of this manipulation. Concord Consortium Ecosystem Task Concord Consortium Osmosis Task	
References	diSessa, A. A. (1993). Toward an 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.	

The "Using" and "Updating, Elaborating, & Correcting Knowledge Representations" *design patterns* are heavily linked with each other; the choice between the two really depends on whether the external data is contradictory or confirmatory with learners' existing mental model. Of course, this information about the match or mismatch (i.e., adequacy or inadequacy) between learners' current mental model and the real world data may be unavailable to the assessor prior to learners' completion of the task. In this case, we would assume coherence (i.e., "Using" *design pattern*) and accommodate accordingly (i.e., "Updating, Elaborating, & Correcting" *design pattern*) if initial results indicate otherwise.

Because there is substantial overlap between the two *design patterns*, the discussion of attributes will not be repeated here. Instead, we will highlight a few additional examples that will help to further instantiate this *design pattern*. It is important to keep in mind that task examples from this *design pattern* can be used to illustrate the previous *design pattern* — and likewise, task examples for the previous *design pattern* can be used to illustrate this *design pattern*.

Below is an assessment simulation item developed by the Concord Consortium, designed to assess learners' understanding about diffusion and osmosis. This simulation is multifaceted, first investigating basic principles of diffusion (i.e., concentration) and then more specialized notions of osmosis through semi-permeable membranes. In the first part of the task, the scenario is framed in terms of the diffusion of a gas into the air, as in the case of a perfume spill in one corner of the room. Learners are shown a particles distribution screen, where white balls represent air and pink balls represent perfume (see Figure 4). They are asked questions regarding the scenario, such as what will happen to the perfume molecules and what are the chances that the perfume molecules will move from one corner of the room to another. These probing questions are asked prior to running the simulation model.

The second part of the task changes the situation slightly, requiring more of the learner in terms of their knowledge. The simulation model shows a cell with a semi-permeable membrane immersed into salt water; the salt molecules are shown in green and the water molecules in white (see Figure 5). Learners are asked to hypothesize about and predict the direction of movement of molecules before running the model. If learners' predictions (and, indirectly, their mental models) are supported by the data, then this would be an example of the "Using Knowledge Representations" *design pattern*. If, however, learners' existing mental models do not include a clause about semi-permeable membranes (i.e., that they allow only certain substances through), then they will expect a similar diffusion process to occur with the salt molecules (as did with the perfume) and may not expect the water molecules to diffuse out of the cell. In this case, they will see a need to correct, elaborate, and/or update their mental models after gathering data from the simulation model; this would be an example of the "Updating, Elaborating, and Correcting Knowledge Representation" *design pattern*.

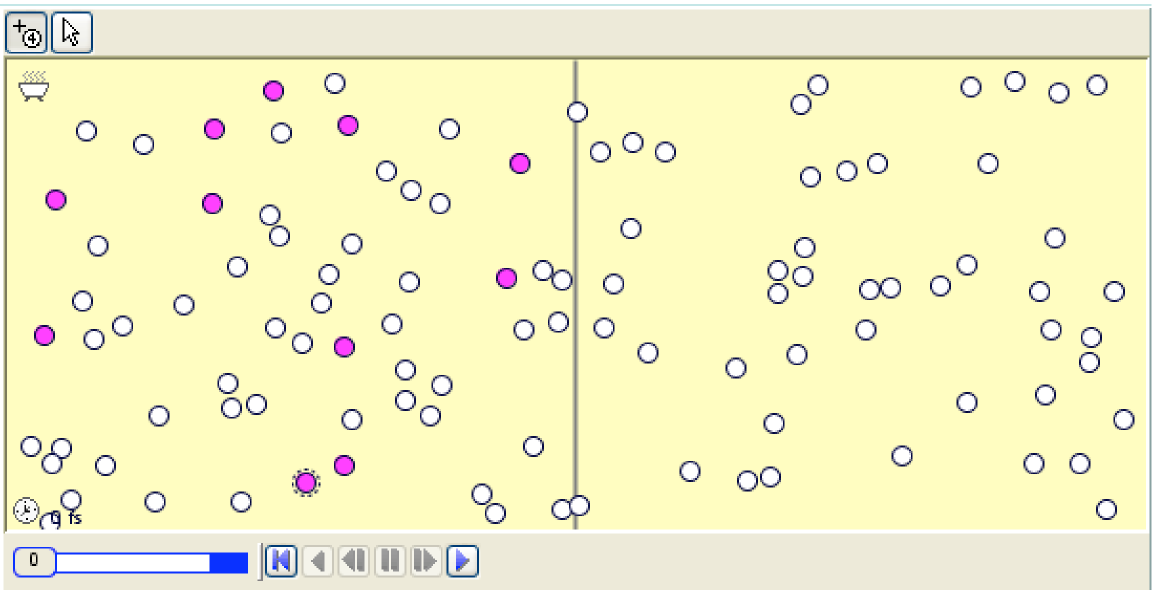

Figure 4. Diffusion of Perfume into the Air

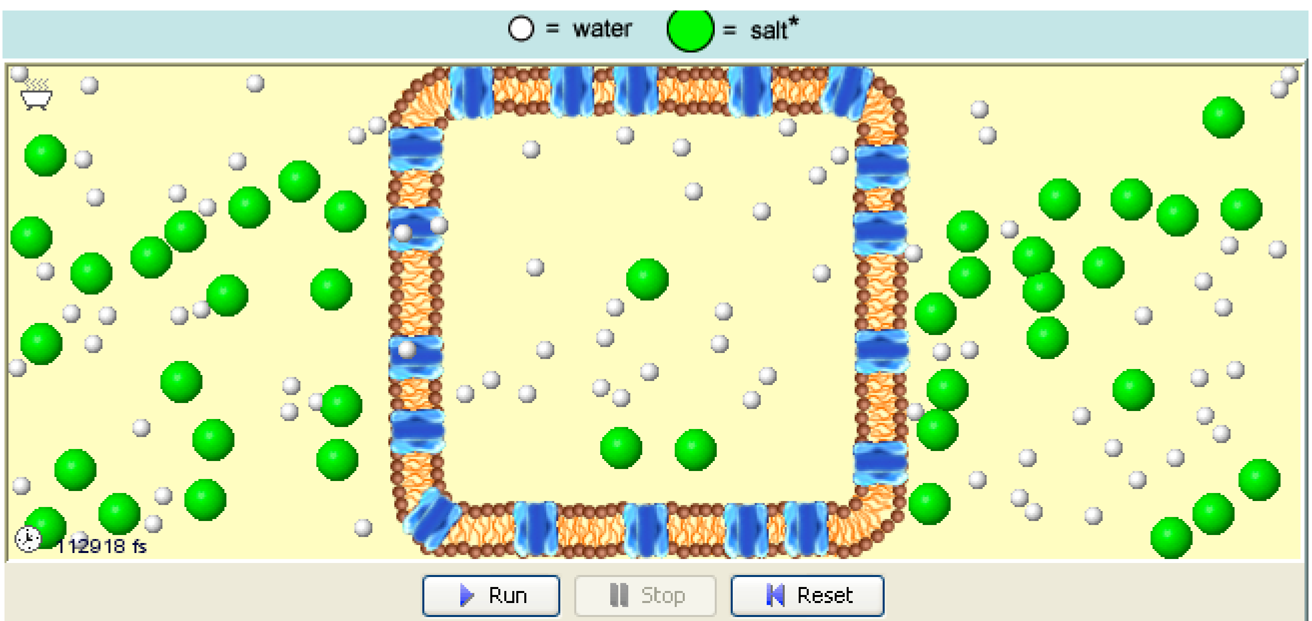

Figure 5. Cell in Salt Water: Semi-Permeable Membranes

The diffusion and osmosis examples demonstrate both the “Using Knowledge Representations” *design pattern* (i.e., first part of the task) and the “Updating, Elaborating, and Correcting Knowledge Representations” *design pattern* (i.e., second part of the task).

5.4 Design Pattern for “Recognizing and Producing Equivalence of Meaning”

Table 7 summarizes a *design pattern* to guide the construction of tasks that aim to assess learners’ ability to recognize and produce equivalence of meaning — taken as an indication of their depth of understanding and indirectly, the intricateness of their mental models.

Summary & Rationale: Multiple representations of concepts — that is, the ability to express the same content, meaning, or knowledge in various ways while at the same time preserving the underlying commonality of meaning between all different expressions — play a key symmetrical role in the *construction* or *augmentation* of mental models. In the “input mode” (i.e., teaching), multiple representations facilitate the *construction* of a good mental model. In the “output mode” (i.e., assessment), a good mental model facilitates the generation of multiple representations. The ability to mentally represent a given meaning in a variety of ways is a prerequisite for, and a marker of, having deep comprehension of conceptual content.

In this *design pattern*, depth of comprehension of newly acquired concepts is assessed by evaluating the learner’s ability to recognize and/or produce multiple representations of content that share a common underlying “meaning” within and across sign-systems. While this *design pattern* is particularly useful for assessing student learning of *newly acquired concepts*, it is also applicable to situations involving the assessment of general knowledge and understanding.

Table 7. Design Pattern for Recognizing and Producing Equivalence of Meaning

ATTRIBUTE	VALUE (S)	COMMENT (S)
Title	Recognizing and Producing Equivalence of Meaning.	
Summary	In this design pattern, depth of comprehension of newly acquired concepts is assessed by evaluating the learners' ability to recognize and/or produce multiple representations of content that share a common underlying 'meaning' within and across sign-systems.	
Rationale	Multiple representations of concepts -- that is, the ability to express the same content, meaning, or knowledge in various ways while at the same time preserving the underlying commonality of meaning between all different expressions-- play a key symmetrical role in the construction or *augmentation* of mental models. In the 'input mode', multiple representations facilitate the *construction* of a good mental model and in the 'output mode,' a good mental model facilitates the generation of multiple representations. The ability to mentally represent a given meaning in a variety of ways is a prerequisite for, and a marker of, having deep comprehension of conceptual content.	
Focal Knowledge, Skills, & Abilities	Ability to decipher the underlying 'meaning' of a *marked* target, consequently being able to recognize or produce alternative expressions of the same meaning. Ability to identify the 'commonality of meaning' amongst a set of alternative expressions, when the target statement is *unmarked* (and hence, must be mentally constructed by the learner). Ability to express or represent a particular 'meaning' in a variety of similar or different looking ways. Ability to transfer (near or far) 'meaning' between or within sign systems. Ability to recognize subtle or obvious differences between the various representations which may or may not alter 'meaning'.	Meaning comprehension occurs when information extracted from the external representation and represented internally via the text surface representation combines and gets linked, associated, and connected with prior knowledge and experiences (Kintsch, 1998).

	Substantive and deep conceptual knowledge about the intricate network of hierarchically connected, organized ideas, relations, and interrelations between various levels of the target concept and its sub-concepts.	
Additional Knowledge, Skills, & Abilities	Sign-system specific and generalized abilities for interpreting and expressing meaning. For instance, the ability to read and interpret graphs and other data knowledge representations. Underlying knowledge activation and retrieval; memory and information processing. Efficiency in comparison and equating of various representations. Ability to manage, monitor, and reflect upon thinking during meaning comprehension and to allocate resources efficiently. Familiarity with type of task	
Potential Observation(s)	1. Pattern of student *selection* (and *non-selection*) of different statement types (T, Q1-Q4) to be meaning equivalent or not (accuracy or partial accuracy): Selecting Q1 statements (i.e., which bear strong surface resemblance as well as meaning equivalence with a target statement) to either mean the same or mean something different than a target. Selecting Q2 statements (i.e., which look different from the target, yet share equivalence-of-meaning with it) to either mean the same as a target or mean something different. Selecting Q3 statements (i.e., which look similar to a target but fail to share its meaning) to either mean the same as a target or mean something different from a target. Selecting Q4 statements (i.e., which neither look similar to the target nor share its meaning) to either mean the same as a target or mean something different from a target. 2. Student *selection* (or *non-selection*) of	In Meaning Equivalence, comprehension deficits are indicated by consistent errors in Q2 and Q3. Interpretation differs for each of the two kinds of errors. A depressed Q2 proportional score is taken to reflect an *over-exclusive* understanding, whereby a learner overlooks or discards meaning equivalent statements because of superficial or surface differences. Low Q3 proportional scores are thought to indicate an *over-inclusive* understanding, whereby a learner wrongly construes non-meaning equivalent statements to be meaning-equivalent based on their surface similarity to the target.

	particular *combination of statement* types to be meaning equivalent or not (accuracy or partial accuracy): Recognizing Q1 as meaning equivalent but failing to recognize Q2 as meaning equivalent because it looks different from a target. Recognizing Q1 as meaning equivalent and also misleadingly choosing Q3 to mean the same as a target because of superficial similarities to a target. 3. Quality (relevancy, accuracy, correctness) of *constructed* representations, *concept statements*, or *verbal defense* of choices.	
Potential Work Product(s)	1. Student selection of meaning equivalent statements from a set of alternatives with a *marked or unmarked* target statement (selection is denoted by shaded or un-shaded bubble sheet or binary scoring of 1 or 0). 2. Student's construction and written expression of the *concept statement* preserved across meaning equivalent statements within each item (i.e., what is the conceptual meaning being preserved between chosen statements?). 3. Student construction/production of alternative meaning equivalent statements (using same or different sign systems) from a *marked* target statement. 4. Verbal discussion and defense of choices for statements preserving equivalence-of-meaning with a small group of colleagues or large class.	Pattern of student selection is important for providing evidence about focal knowledge, skills, and/or abilities. A concept statement details a particular idea (e.g., particular aspect of a concept etc.,) or group of ideas that the domain expert or instructor deems pertinent to the course.
Potential Rubric(s)	Five types of scores, at the item-specific level as well as across items, may be used for providing feedback to the learner or instructor following test completion: (a) positive and negative partial scores; (b) quadrant-specific partial scores; (c) global (GL) score;	The scoring system is sufficiently detailed to offer a multitude of scoring options, each revealing a separate element of the learning process with a different intended purpose whether for evaluative assessment, formative assessment, remediation, or modification to the curriculum. Some scores such as the Global

		Score, provides an overview of student performance and understanding. Other scores such as the quadrant-specific partial scores allow a more detailed analysis of students' strength and weakness in preparation for detailed feedback and subsequent remediation efforts.
Characteristic Feature(s)	Tasks that ask student to work with or through multiple representations of knowledge concepts.	
Variable Feature(s)	Whether to ask students to *select* or *construct* multiple representations of meaning (type of response required). Whether *the target statement* should be *marked* or *unmarked.* Varying the *content, scope, and nature* of the representation of the concept being probed in the target statement (e.g., inclusion of concepts with conditional meaning). *Working memory* and *cognitive processing load* required. Whether to have *multiple* sign-systems within the same item (e.g., text, algebraic, images). Whether to ask for meaning transfer between or within sign-systems (i.e., near or far transfer). Whether to minimize the influence of *external factors* such as complexity of language for certain populations (e.g. reading disabled), when language is not the focus of the assessment. Whether to administer the test via *online or paper-based* environment. Whether to use the assessment as a self scoring *learning tool* or formal *assessment tool.* Whether to maximize or limit the *level/complexity/ range (and integration)* of	

	conceptual knowledge required to solve problem. Variables specific to the ME approach: Whether to include a certain combination of other *meaning equivalent statement types* (e.g., more Q2s than Q1s make the item more difficult) within an item. Whether to include a certain combination of other *non-meaning equivalent statement types* (e.g., more Q3s than Q4s) within an item. Whether to have *2, 3, 4, or 5* meaning equivalent statements included within an item and the *nature of their representation*.	
Exemplar Tasks	Refer to the various Meaning Equivalence Test examples.	
References	Sigel, I. E. (1999). Approaches to representation as a psychological construct: A treatise in diversity. In I. E. Sigel (Ed.), Development of mental representation: Theories and applications (pp. 3-12). Mahwah, NJ: Erlbaum. Shafrir, U. (1999). Representational competence. In I. E. Sigel (Ed.), The development of mental representation: Theory and applications (pp. 371-389). Mahwah, NJ: Erlbaum. Seeratan, K. (2006) [1]. Assessing and enhancing learning outcomes in an architectural context: Meaning equivalence methodology versus traditional formats of testing. Ph.D. Dissertation.	

Focal KSAs: According to Kintsch (1998), meaning comprehension occurs when information extracted from the external representation and represented internally via the text surface representation, combines and gets linked, associated, and connected with prior knowledge and experiences. If the mental model constructed is solid, then learners should be able to a) *construct or represent* a particular meaning in a variety of similar or different looking ways; b) *recognize* equivalence of meaning among a set of given alternative expressions; and c) identify the commonality of meaning being preserved in a) and b). These are the Focal KSAs of the "Recognizing and Producing Equivalence of Meaning" (henceforth to be called the "Meaning Equivalence (ME)") *design pattern*.

Additional KSAs are other skills that may be required in a task developed from this *design pattern*. An assessment designer must consider, for instance, sign–system specific and generalized abilities for interpreting and expressing meaning. This includes being able to read text or interpret/construct graphs and/or other data representations. Another consideration is learners' familiarity with the task type. Some tasks, such as writing essays or completing multiple-choice items, have been used in traditional testing situations for centuries and so are relatively familiar to students. Although the "Meaning Equivalence (ME)" format has some superficial similarities with the multiple-choice (MC) format, the cognitive skills that are required for solving each type are divergent. Like MC, in ME, students are given a set of five alternatives from which they must choose the most appropriate answer(s). In MC, a "stem" is provided, and there is usually only one correct "leaf," which the learner must choose from the set of given alternatives. In ME, the "stem" may or may not be presented, and the "leaf" may be two or more of the presented alternatives. While both formats are considered selected-response types, ME also has a constructed-response component (i.e., similar to Essay) where the learner is required to construct a verbal or written description of the meaning being preserved by the statements selected as meaning equivalent. It is expected (and has been found) that learners initially have difficulty with the unfamiliarity of the ME task type, but this difficulty becomes negligible after some exposure to the format. Hence, familiarity with the task type is an Additional KSA — a construct irrelevant source of differential performance or difficulty that may affect student performance on this task. Other Additional KSAs for this *design pattern* include learners' knowledge, practice, or familiarity with comparing or equating various representations; and their ability to manage, monitor, and reflect upon thinking during meaning comprehension and to allocate mental resources efficiently.

Variable Task Features are characteristics of the task that the assessment designer can choose to vary to shift the difficulty or focus of tasks. For the ME *design pattern*, one can choose to vary, among other things, whether to ask learners to *select* or *construct* multiple representations of meaning (i.e., type of response required). Response procedures tap varying cognitive skill sets. For instance, the constructed-response format require students to generate information to answer a question without access to any retrieval cues from potential alternatives and is typically deemed to be tapping higher level cognitive skills. The selected-response format, on the other hand, allows students to choose the best answer from a set of alternatives provided and is typically deemed to be tapping relatively lower level cognitive skills (although this is not necessarily the case).

In the Representational Competence of Algebra test (i.e., RC-Alg), designed to assess learners' ability to recognize and produce expressions conveying equivalent algebraic meanings across different sign systems, students were given an algebraic equation such as, $3x + 5 = 14$, and asked to *produce* a written story that represents this equation.

The Representational Competence of Text test (i.e., RC-Text), designed to assess deep comprehension of English text at the sentence level, is an example of a test (i.e., set of tasks) that uses the ME *design pattern* to develop items that have a selected-response format. Students are given a set of five statements and are asked to mark all statements which preserve the same underlying meaning. An example of an item from RC-Text test is shown below. Note that, unlike the example below, the statement types (e.g., Q1, Q2, Q3, Q4, & Target) were not labeled in the actual test.

A *It was chosen by the critics as the best movie of the year.* [**Q1**]
B *The best movie of the year was met with critical acclaim.* [**Q3**]
C *The title, "The best movie of the year," was bestowed on it by the critics.* [**Q2**]
D *The critics acclaimed it "the best movie of the year."* [**Target**]
E *No doubt it was the best movie of the year.* [**Q4**]

In the ME approach, there are four different statement types: Q1 statements bear strong surface resemblance as well as meaning equivalence with a target statement; Q2 statements look different from a target, yet share equivalence-of-meaning with it; Q3 statements look similar to a target but fail to share its meaning; while Q4 statements neither look similar to a target or share its meaning. Learners are expected to select statements that mean the same as a target (i.e., Q1, Q2) but not select statements that mean something different from a target (i.e., Q3, Q4). In the RC-Text example above, learners would be expected to select options A, C, and D as meaning equivalent. Under this approach, comprehension deficits are indicated by consistent errors in failing to select Q2 while selecting Q3 (Shafrir, 2001; Seeratan, 2006[1]).

Another Variable Feature for this *design pattern* pertains to the target statement — specifically, whether it should be *marked* or *unmarked.* If the target statement is unmarked, as in the case of the selected-response scenario (see RC-Text example), learners are unaware of the "stem" of the problem (i.e. the target meaning that they must use as the basis for determining equivalence of meaning among the given alternatives). Therefore, in trying to decipher meaning equivalence among alternative statements, the common meaning must first be mentally constructed by learners. Learners begin by looking for equivalence of meaning across any two of the five statements in the item (e.g., compare A to B then A to C, then A to D, then A to E, then B to C, B to D, etc). Once the underlying meaning equivalence has been detected between two statements, learners then test this meaning against the meaning of each of the three remaining statements. On the other hand, if the target meaning is explicitly marked, then the task is comparatively easier; learners merely have to use this meaning as the "stem" in their search for alternatives that transmit the same meaning. Note that although having an unmarked target allows less scaffolding and assesses a higher level of processing, it places high cognitive load demands on the learner. The marked target is one method devised for reducing cognitive load demands of the task (Shafrir, 2001; Seeratan, 2006[1]).

In addition, for this *design pattern*, one can also choose to vary the number of sign-systems used within items. For instance, items may require movement within *one* sign-system as in the RC-Text example (i.e., from text to text) or they can require movement between *multiple* sign-systems as in the RC-Algebra item (e.g., from algebraic equation to text). As discussed earlier, the cognitive mechanisms involved in meaning extraction has been shown to vary both within and between sign-systems or modalities. Refer to Seeratan (2006) [1] for more details.

Other Variable Features include varying the *content, scope, and nature* of the representation of the concept being probed in the target statement (e.g., inclusion of concepts with conditional meaning). Here, content can be at general or specific levels in any domain. Another option is varying w*orking memory* and *cognitive processing load* required. One way that this can be varied is by deciding to explicitly mark the target statements. Another way is by minimizing, as much as possible, the language and/or processing demands of each representation. Finally, depending on the purpose of intended usage, a designer may choose to use this task as an instructional tool for helping learners to build good mental models;

alternatively, the task may be used as an assessment tool for gauging the organization, elaborations, and hierarchical structure of the learners' current intact mental models (Seeratan, 2006[1, 2]).

Variable Features specific to the ME approach include: whether to include a certain combination of other *meaning equivalent statement types* (e.g., more Q2 than Q1 types make the item more difficult) within an item; whether to include a certain combination of other *non-meaning equivalent statement types* (e.g., more Q3s than Q4s makes the item more difficult) within an item; whether to have *2, 3, 4, or 5* meaning equivalent statements included within an item and the *nature of their representation* (Shafrir, 2001; Seeratan , 2006[1]).

Work Products that can be called for in ME tasks include information about:

1. Learners' selection of ME statements from a set of alternatives, denoted by shaded or un-shaded bubble sheet or binary scoring of 1 or 0 — that is, which statements, or combination thereof, did learners choose as preserving the same meaning as a marked target or unmarked target? The pattern of selection is important for providing evidence about the Focal KSAs.

2. Learners' mental construction and written expression of the *concept statement* being preserved across ME statements within each item – here, learners are asked to formulate and provide a written statement of the meaning that is being preserved and which they themselves have constructed as in the case of an unmarked target.

3. Learners' construction/production of alternative ME statements (using the same or different sign-systems) from a *marked* target statement. This is relevant to constructed-response ME formats.

4. Verbal discussion and defense of choices for statements preserving equivalence-of-meaning with a small group of colleagues or large class. Here, learners get together with classmates to

debate over their choice of ME statements; such a debate would act to enhance learning, facilitate transfer, and help to clarify for each learner what they did and did not understand (Seeratan, 2006[1]).

Several kinds of Potential Observations can be evaluated from these Work Products:

1. The pattern and accuracy (whole or partial) of student *selection* (and *non-selection*) of different statement types.

2. The pattern of student *selection* (or *non-selection*) of particular *combinations of statement* types to be meaning equivalent or not (accuracy or partial accuracy): for instance, recognizing Q1 as meaning equivalent relative to a target but failing to recognize Q2 as meaning equivalent because it looks different from a target; or recognizing Q1 as meaning equivalent and also misleadingly choosing Q3 to mean the same as a target because of superficial similarities.

3. The quality (relevancy, accuracy, correctness) of *constructed* representations (as in constructed-response format), *concept statements* (i.e., what is the meaning being preserved between/among statements), or *verbal defense* of choices (Seeratan, 2006[1]).

Potential Rubrics: The ME scoring system is sufficiently detailed to offer a multitude of scoring options, each revealing a separate element of the learning process with a different intended purpose, whether for evaluative assessment, formative assessment, remediation, or modification to the curriculum. The ME system has scoring options at the item-specific level as well as across items. Below are brief descriptions of three of the ME scoring options, which may be used for providing feedback to the learner or instructor following test completion:

a) *Partial Score* is a refined score. There are two partial scores: positive and negative. Positive partial score is the proportion of correct responses of those statements in each item that should have been and were, in fact, marked for equivalence-of-meaning (i.e.

Target, Q1, and Q2). Negative partial proportional score is the proportion of correct responses of those statements in each item that should not, and have not been marked for equivalence-of-meaning (i.e. Q3 and Q4). Partial scores can be calculated at both the item and test levels.

b) *Quadrant-Specific Score* reflects the learners' proportional score on each of the statement types (Target, Q1 to Q4) included within a particular item and between items. Quadrant-specific scores can be calculated at both the item and test levels.

c) *Global Score (GL)* is simply the score in terms of proportion correct. In scoring each item for global feedback, the learner receives a score of 1 for each item answered correctly. In this case, the operational definition of "answered correctly" is that all five statements must be correctly checked (if Target, Q1 or Q2) or unchecked (if Q3 or Q4) by the learner; that is, he/she must check all statements that share equivalence-of-meaning with a target (i.e., Q1 and Q2) and must fail to check all statements that do not share such equivalence-of-meaning (i.e., Q3 and Q4).

Some scores, such as the Global Score, provide an overview of learners' performance and understanding. Other scores, such as the quadrant-specific partial scores, allow a more detailed analysis of learners' strength and weakness in preparation for detailed feedback and subsequent remediation efforts (Shafrir, 2001; Seeratan, 2006[1]).

5.5 Design Pattern for "Re-Expressing Information"

Table 8 summarizes a *design pattern* to guide the construction of tasks that aim to assess learners' ability to re-express information — that is, to take one data representation and transform it into another while preserving its' meaning. Note that this *design pattern* is very similar to the “Recognizing and Producing Equivalence of Meaning” *design pattern* just discussed. In fact, it can be seen as an instantiation of it, with a narrower focus on *data representation* — specifically, particular forms of external knowledge representations or inscriptions (Lewandowsky & Behrens, 1999).

Summary & Rationale: In this *design pattern*, learners encounter data organized in one (or more) representational forms and are asked to re-express it in terms of one (or more) different representational form(s). A representational form is a schema for organizing information; it has conventions such that spatial or relational relationships of elements in the representation correspond to relationships among entities, processes, or events in the real world. Re-expressing data involves recognizing the elements being addressed in a particular representation, understanding the relationships amongst them, then producing/identifying/critiquing the mapping of those relationships unto a different form. Representational forms for data can be general (e.g., charts, graphs, and tables) and/or specialized.

Table 8. Design Pattern for Re-Expressing Information

ATTRIBUTE	VALUE (S)	COMMENT (S)
Title	Re-expressing data.	This design pattern is similar to the Meaning Equivalence design pattern but is more specifically focused on *data representation – s*pecifically, particular forms of external knowledge representations, or inscriptions (Lewandowsky & Behrens, 1999).
Summary	In this design pattern, a learner encounters data organized in one or more representational forms and is asked to re-express it in terms of one or more different representational form (s).	Representational forms for data can be general (e.g., charts, graphs, and tables) and/or specialized. A representational form is a schema for organizing information; it has conventions such that spatial or relational relationships of elements in the representation correspond to relationships among entities, processes, or events in the real world. Re-expressing data involves recognizing the elements being addressed in a particular representation, understanding the relationships among them, then producing/ identifying/ critiquing the mapping of those relationships unto a different form.
Rationale	Multiple representations -- that is, the ability to express the same content, meaning, or knowledge in various ways while at the same time preserving the underlying commonality of meaning between all different expressions-- play a key symmetrical role in the construction or *augmentation* of mental models. In the 'input mode', multiple representations facilitate the *construction* of a good mental model and in the 'output mode,' a good mental model facilitates the generation of multiple representations. The ability to	Scientific data are measurements, observations, counts, or classifications of real-world phenomena, organized in terms of some scientific representational form. They may be organized in a standard way, or in a way connected by a particular scientific understanding of the situation at hand.

	mentally represent a given meaning in a variety of ways is a prerequisite for, and a marker of, having deep comprehension of conceptual content.	
Focal Knowledge, Skills, & Abilities	Ability to decipher the underlying 'meaning' (including relations and interrelations between elements) of data represented in the target representational form (s). Ability to interpret and integrate data meaning from multiple representational forms. Ability to re-express data while preserving underlying meaning, relationships, and interrelations between elements from one or more representational forms. Substantive and deep conceptual knowledge about the intricate network of hierarchically connected, organized ideas, relations, and interrelations between various levels of the target concept and its sub-concepts.	
Additional Knowledge, Skills, & Abilities	Knowledge of different representational forms for data expression (how to interpret and re-express meaning through them). Underlying knowledge activation and retrieval; memory and information processing. Relevant domain knowledge. Knowledge of data interpretation, equating, computation, manipulation (e.g., knowledge of mathematics) Efficiency in comparing, evaluating, and equating multiple representations. Ability to manage, monitor, and reflect upon thinking during meaning comprehension and meaning making process and to allocate mental resources efficiently.	
Potential Observation(s)	Combining data from multiple representational forms into a new representational form. Constructing new representation with appropriate design, layout, units, and terminologies etc.,	

	Critique or rationale for other learners' re-expressions of the given data. Explanation or rationale for learners' own re-expressions of the data. Identification of correct/incorrect re-expressions from given ones. Identification of appropriate representational forms for given data. Putting data into new representation correctly. Quality (relevancy, accuracy, correctness) of re-expressions, explanations of, and rationale for choice of format.	
Potential Work Product(s)	Learners' construction or re-expression of data in new 'meaning preserving' format (s). Learners' construction and written expression of the data relationships and interrelationships being preserved across different meaning equivalent representational forms. Identification, from given possibilities, of most appropriate rationale for using a particular representational form. Explanation of appropriate/inappropriate representational forms (verbal/written response) for given data. Verbal discussion and defense of choice for representational form with a small group of colleagues or large class.	
Potential Rubric(s)	http://www.globeassessment.sri.com/genericrubric.htm	
Characteristic Feature(s)	Tasks that ask learners to work with or through multiple representational forms for data expression; Learners must recognize elements of importance in a particular target representational form (i.e., elements and relationships) and then re-express the elements in different 'meaning preserving' forms.	One or more representational forms are required for original presentation of data; one or more different representational forms are involved in the re-expression.
Variable Feature(s)	Whether to ask learners to *select* or *construct* different representational forms encoding the same data meaning (type of response	

	required). Varying the *content, scope, and nature* of data being probed in the target representational form. *Working memory* and *cognitive processing load* required. Complexity of the target representational form (e.g., complex relationships, numbers of variables, etc.,) Whether to use *multiple* sign-systems (e.g., text, algebraic formulas, tables, graphs) for representational forms. Number of representational forms that one must integrate and/or create (e.g., combining information from *one or more* target representation(s) to form *one or more* new representation (s)). Familiarity with different data representation forms (e.g., algebraic). Directness of translation (e.g., straightforward or involving computation or transformation of information). Whether to maximize or limit the *level/complexity/ range (and integration)* of conceptual knowledge required to address the problem.	
Exemplar Tasks	Refer to GLOBE Activities. Refer to Meaning Equivalence Activities.	
References	Sigel, I. E. (1999). Approaches to representation as a psychological construct: A treatise in diversity. In I. E. Sigel (Ed.), Development of mental representation: Theories and applications (pp. 3-12). Mahwah, NJ: Erlbaum. Seeratan, K. (2006) [1]. Assessing and enhancing learning outcomes in an architectural context: Meaning equivalence methodology versus traditional formats of testing. Ph.D. Dissertation.	

Because this *design pattern* is an instantiation of the one before it, the previous discussion will not be repeated, although it is applicable here as well. Only attributes specific to this *design pattern* will be discussed.

The Focal KSAs for this *design pattern* include: deciphering the underlying "meaning" of data represented in the target representational form(s) and includes the: ability to recognize the elements being addressed in the representational form, how these elements are hierarchically connected, organized, their relations, and interrelations; ability to interpret and integrate data meaning from multiple representational forms; ability to re-express data while preserving underlying meaning, relationships, and interrelations between elements from one or more representational forms.

Additional KSAs: Specific to this *design pattern*, an assessment designer must consider the student's familiarity or knowledge of different representational forms for data expression (i.e., how to interpret and re-express meaning through them) and knowledge of data interpretation, equating, evaluating, computation, manipulation (e.g., knowledge of mathematics).

Variable Task Features: Specific to the re-expressing *design pattern*, one can choose to vary, among other things: complexity of the given and target representational form (e.g., complexity of relationships, numbers of variables, number of sign-systems used, whether the translation is within or between sign-systems etc.); number of representational forms that one must integrate and/or create (e.g., combining information from *one or more* target representation(s) to form *one or more* new representation (s)); familiarity with different data representation forms (e.g., algebraic); and, directness of translation (e.g., straightforward or involving computation or transformation of information).

Work Products that can be called for in the re-expressing tasks include information about: learners' construction or re-expression of data in new "meaning preserving" format (s); learners' construction and written expression of the data relationships and interrelationships being preserved across different meaning equivalent representational forms; identification, from given possibilities, of the most appropriate rationale for using a particular representational form; explanation of appropriate/inappropriate representational forms (verbal/written response) for given data; and verbal discussion and defense of choice for representational form with a small group of colleagues or a large class.

Several kinds of Potential Observations can be evaluated from these Work Products. These include the accuracy and quality of: combining data from multiple representational forms into new representational form(s); constructing new representation using appropriate representational form, design, layout, units, and terminologies; critique of other learners' re-expressions of the given data; and learners' explanation or rationale for their own re-expressions of the data.

Instantiations of this *design pattern* can be found in ME tasks. Although this *design pattern* is more specific to data re-expression than the *Recognizing and Producing Meaning Equivalence design pattern*, it is general enough to be effectively applied across many content areas, and at different ability levels. An example of a ME task can be taken from the ME Business test where students are given data in table format and asked to re-express its meaning in the form of a graph. The Globe assessment tasks also provide several instantiations of this *design pattern*. For instance, in the "plot of four planets" task, a National Assessment of Educational Progress (NAEP) grade 12 item (see Mislevy et al., 2003), students are given a simple table of data about four planets, their periods of revolution about the sun, and rotation about their axes. They are given explicit instruction to translate the data table into a line graph, on a given axes and graph structure. Even more specifically, they are told to “plot a point for each of the four planets showing the planet’s period of revolution, and its mean distance from the sun. Then draw the line or curve that best illustrates the relationship between the period of revolution and the mean distance from the sun that is suggested by the points.” This is a very simple, highly scaffolded item; the data are very simple, involving the relationship between only two variables and just a few data points. There is some extraneous data in the table that students are supposed to know they must ignore, this adds a bit of difficulty to the task.

In a more difficult example — a task from the Maryland School Performance Assessment Program (MSPAP) — grade 8 students work in groups, and their task is to take the information from two representational forms (i.e., a table and orbit diagram) and re-represent the information in a new form called an “orbit data log.” This task is similar to the NAEP task, but it involves integrating two representational forms and has substantially less cueing. In addition, the new representational form that they are asked to create is unfamiliar, which adds another dimension of complexity (Mislevy et al., 2003).

In another simple example, from the BioKIDS assessments, students are given data in table form (e.g., observations of the number of animals in different parts of the school yard) and asked to choose from a set of four graphs the one that best represents the data given in the table. Another assessment example from the Constructing Data Modeling Worlds (CDMW) project gives students a scenario in text form and asks them to choose which one among four graphs best represents the relationship outlined in the text.

6.0 Conclusion

Fundamental to both pedagogical and assessment approaches is an understanding of how students learn, specifically how they represent, organize, structure, manipulate, retrieve, and convey informational content gathered from the external realm. Knowledge about the intricate structure and function of our mental models can provide pertinent insights into the quality of knowledge thus far attained, including whether it is organized for effective action and learning, accordingly suggesting instructional pathways for facilitating future knowledge acquisition. Having appropriate valid and reliable methods for assessing learners' knowledge representations is an integral part of the learning process.

In this report, we presented an overview of dimensions essential to internal representations of knowledge and discussed issues surrounding how that knowledge can be assessed, accenting in particular how *design patterns* — conceptual tools used for providing well thought out navigation in the design of inquiry tasks, coordination of and provision of tools to facilitate them — can serve to improve our capacity and efficiency for doing so.

Gathering research insights from multidisciplinary fields, including cognition and measurement, *design patterns* help to transform research-based knowledge about assessments design into organized templates, accessible and easy to use in educational practice; consequently, *design patterns* act as agents in helping to facilitate and close the gaps within research to practice pursuits.

REFERENCES

Ainsworth, S. E. (1999). A functional taxonomy of multiple representations. *Computers and Education, 33*, 131-152.

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, memory, & cognition, 19(5),* pp. 1061-1070.

Anderson, J. R., & Bower, G. H. (1973). *Human Associative Memory.* Winston-Wiley Publishers.

Anderson, J. R. (1990). *Cognitive Psychology and Its Implications*. New York: Freeman and Company.

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85,* 249-277.

Anderson, J. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Bainbridge, L. (1979). Mental models and cognitive skill: the example of industrial process operation. In Y. Rogers, A. Rutherford, and P.A. Bibby (Eds.), *Models in the Mind: Theory, Perspective, and Application.* Academic Press, pp. 119-512.

Boose, J. H. (1986). *Expertise Transfer for Expert System Design.* Amsterdam: Elsevier**.**

Bransford, J. D., Brown, A. L., & Cocking, R. (1999). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.

Carley, K., & Palmquist, M. (1992). Extracting, representing, and analyzing mental models. *Social Forces, 70(3),* pp. 601-636.

Carley, K. (1984). *Constructing Consensus.* Ph.D. dissertation. Harvard University.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5(2),* pp.121-152.

Clancey, W. J. (1993). Situated action: A neuropsychological interpretation response to Vera and Simon. *Cognitive Science, 17,* 87-116.

Clarke, B. (1987). Knowledge acquisition for real-time knowledge-based systems. *Proceedings of the First European Workshop on Knowledge Acquisition for Knowledge Based Systems,* Reading University, UK, 2-3 September.

Collins, A., & Gentner, D. (1987). How people construct mental models. In D. Holland & N. Quinn (eds.), *Cultural models in thought and language* (pp. 243-265). Cambridge: Cambridge University Press.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies, 41(6),* 801-849.

Cordingley, E. S. (1989). Knowledge elicitation techniques for knowledge-based systems. In D. Diaper (ed.), *Knowledge Elicitation: Principles, Techniques, and Applications*. John Wiley & Sons, pp. 89-175.

Craik, K. (1943). *The Nature of Explanation*. Cambridge University Press, UK: Cambridge.

DeBarger, A. H., & Riconscente, M. (2005). *An example-based exploration of design patterns in measurement (*PADI Technical Report 8). Menlo Park, CA: SRI International.

Deloache, J. S. (1989). The development of representation in young children, *Advances in Child Development and Behavior, 22,* pp. 1-40.

Dienes, Z. (1973). *The six stages in the process of learning mathematics*. Slough, England: NFER-Nelson.

diSessa, A. (1983). Phenomenology and the evolution of intuition. In D. Gentner & A.L. Stevens (Eds.), *Mental models* (pp. 15-33). Hillsdale, NJ: Erlbaum.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10 (2 & 3),*105-225.

Dixon, P. (1991). Learning to Operate Complex Devices: Effects of Conceptual and Operational Similarity. *Human Factors, 33 (2),* pp. 103-120.

Dretske, F. (1986). Aspects of cognitive representation. In M. Brand, & R. M. Harnish (Eds.), *The Representation of Knowledge and Beliefs*. Tucson, AZ: The University of Arizona Press.

Dutke, S. (1996). Generic and generative knowledge: Memory schemata in the construction of mental models. In W. Battmann & S. Dutke (Eds.), *Processes of the Molar Regulation of Behavior.* Pabst Science Publishers, pp. 35-54.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. MIT Press, Cambridge, MA: USA.

Evans, A. W., Jentsch, F., Hitt, J. M., Bowers, C., & Salas, E. (2001). Mental Model Assessments: Is there convergence among different methods? *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting,* pp. 293-296.

Fincher, S., & Tenenberg, J. (2005). Making sense of card sorting data. *Expert Systems, 22(3),* 89-93.

Finke, R. A. (1989). *Principles of mental imagery*. Cambridge: MIT Press.

Fiske, S. T., & Taylor, S. E. (1984). *Social Cognition*. Addison-Wesley.

Fodor, J. A., (1981). *Representations: Philosophical Essays on the Foundation of Cognitive Science*. Sussex: The Harvester Press.

Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, MA: MIT Press/Bradford.

Garnham, A. (1997). Representing information in mental models. In M. A. Conway *Cognitive Models of Memory: Studies in Cognition*. MIT Press, 149-172.

Geiwitz, J., Kornell, J., & McCloskey, B. (1990). *An Expert System for the Selection of Knowledge Acquisition Techniques.* Technical Report 785-2. Santa Barbara, CA: Anacapa Sciences.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*(3)*,* 430-445.

Gibson, J. J. (1977). The theory of affordances. In R. Shawn and J. Bransford (Eds.) *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Lawrence Erlbaum Associates, pp.67-82.

Gitomer, D. H., & Steinberg, L. S. (1999). Representational issues in assessment design. In I.E. Sigel (Ed.), *Development of mental representation* (351-370). Mahwah, NJ: Erlbaum.

Glaser, R., & Chi, M.T.H. (1988). Overview. In M.T.H. Chi, R. Glaser, & M.J. Far (Eds.), *The Nature of Expertise,* pp.xv-xxviii. Hillsdale, NI: Lawrence Erlbaum.

Glaser, R. (1988). Cognitive science and education. *International Social Science Journal, 40,* 21-44.

Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs. forcing*. Mill Valley, CA: Sociology Press.

Greeno, J. G., & Moore, J. L. (1993). Situativity and symbols: Response to Vera and Simon. *Cognitive Science, 17*, 49-59.

Hall, E. P., Gott, S. P., & Pokorny, R. A. (1995) *A procedural guide to cognitive task analysis: The PARI methodology.* Armstrong Research Laboratory, Brooks AFB, Texas.

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

Heylighen, F. (2001). Bootstrapping knowledge representations: From entailment meshes via semantic nets to learning webs. *Kybernetes, 30,* 691-722.

Hively, W., Patterson, H.L., & Page, S.H. (1968). A "universe-defined" system of arithmetic achievement tests. *Journal of Educational Measurement*, *5*, 275-290.

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.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Inference, and Consciousness.* Cambridge University Press, UK: Cambridge.

Kaput, J. J. (1988). Representation systems and mathematics. In C. Janvier (Ed.), *Problems of Representation in the Teaching and Learning of Mathematics.* Hillsdale, New Jersey: Lawrence Erlbaun Associates.

Kaput, J. J. (1989). Linking representations in the symbol systems of algebra. In S. Wagner & C. Kieran (Eds.), *Research issues in the learning and teaching of algebra* (pp. 167-194). Hillsdale, NJ: Erlbaum.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science, 8(3),* pp.255-273.

Kintsch, W., and Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109-129.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363-394.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review, 95,* 163-182.

Kintsch, W. (1996). Mental representations in cognitive science. In W. Battmann & S. Dutke (Eds.), *Processes of the molar regulation of behavior.* Pabst Science Publishers, pp. 17-33.

Kintsch, W. (1998). The representation of knowledge in minds and machines. *International Journal of Psychology, 33 (6),* 411-420.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Knorr-Cetina, K. (1981). *The Manufacture of Knowledge*. Pergamon.

Kosslyn, S. M. (1978). Imagery and internal representation. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and Categorization* (pp. 217-258). Hillsdale, New Jersey: Lawrence Erlbaum.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2002). Mental imagery doesn't work like that. *Behavioral and Brain Sciences, 25*, 198-200.

Kowalski, B., & Van Lehn, K. (1988). Cirrus: inducing subject models from protocol data. *Proceedings of the 10th Annual Conference of the Cognitive Science Society*, pp. 623-629, Quebec, Canada, 17-19 August.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika, 29*, 115-129.

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

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

Lehrer, R., & Schauble, L. (2002). Symbolic communication in mathematics and science: Co-constituting inscription and thought. In E.D. Amsel & Byrnes (Eds.), *Language, Literacy, and Cognitive Development: The Development and Consequences of Symbolic Communication.* (pp.167-192). Mahwah, NJ: Lawrence Erlbaum Associates.

Markman, A. B. (1999). *Knowledge Representations.* Mahwah, NJ: Erlbaum.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*(1), 1-19.

Mead, G. H. (1964). The genesis of the self and social control. In A.J Reck (Eds.), *Recent American Philisophy: Studies of Ten Representative Thinkers.* Bobbs-Merrill Publishers..

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

Mislevy, R. J., Behrens, J. T., Bennett, R. E., Demark, S. F., Frezzo, D. C., Levy, R., Robinson, D. H., Rutstein, D. W., Shute, V. J., Stanley, K., & Winters, F. I. (2007). *On the Roles of External Knowledge Representations in Assessment Design*, CRESST/CSE Report 722.

Mislevy, R., Hamel, L., Fried, R., G., Gaffney, T., Haertel, G., Hafter, A., Murphy, R., Quellmalz, E., Rosenquist, A., Schank, P., Draney, K., Kennedy, C., Long, K., Wilson, M., Chudowsky, N., Morrison, A., Pena, P., Songer, N., Wenk, A. (2003). *Design* *patterns for assessing science inquiry* (PADI Technical Report 1). Menlo Park, CA: SRI International.

Moray, N. (1999). Mental models in theory and practice. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application. Attention and Performance.* MIT Press, pp.293-297.

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education, 30*(2)*,* 122-147.

Nelissen, J.M., & Tomic, W. (1996). Representation and Cognition. *Eric Journals*

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Ok-Choon, P., & Stuarts, G. S. (1995). Dynamic characteristics of mental models and dynamic visual displays. *Instructional Science, 23(5-6),* pp.303-320.

Olson, J. R., & Rueter, H. H. (1987). Extracting expertise from experts: methods for knowledge acquisition. *Expert Systems, 4(3),* p. 152-168.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B.B Lloyd, (Eds.), *Cognition and Categorization* (pp. 259-303). Hillsdame, New Jersey: Lawrence Erlbaum Associates.

Peirce, C. S. (1906). Prolegomena to an apology for pragmaticism. *The Monist*, *16,* 492-497.

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioural and Brain Sciences, 25,* 157-238.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Transactions on Systems, Man, and Cybernetics, 15(2),* pp.234-243.

Rouse, W.B., & Morris, N.M. (1986). On Looking Into the Black Box: Prospect and Limits in the Search for Mental Models. *Psychological Bulletin, 100(3),* pp. 349-363.

Sackett, G. P. (1977). *Observing Behavior, Volume 1.* Baltimore, MD: University Park Press.

Sasse, M-A. (1991). How to t(r)ap users' mental models. In M. J. Tauber, and D. Ackermann (eds.), *Mental Models and Human-Computer Interaction* 2. Elsevier Science Publishers B. V., pp. 59-79.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2)*,* 141-156.

Schum, D. A. (1994). *The evidential foundation of probabilistic reasoning.* New York: Wiley.

Schvaneveldt, R.W. (1990). *Pathfinder Associative Networks: Studies in Knowledge Organization*. Ablex Publishing, Norwood: NJ.

Seeratan, K. L. (2006)[1]. *Assessing and enhancing learning outcomes in an architectural context: Meaning equivalence methodology versus traditional formats of testing.* Ph.D., Dissertation published by the National Library of Canada.

Seeratan, K. L. (2006)[2]. Using the meaning equivalence method to assess and enhance learning outcomes. *Conference Proceedings of the Inter-service/Industry Training, Simulation, and Education Conference (I/ITSEC)*, Orlando, Florida.

Shadbolt, N., & Burton, M. (1990). Knowledge elicitation. In J. R. Wilson & E.N. Corlett (Eds.), *Evaluation of Human Work: A practical ergonomics methodology.* Taylor and Francis, pp. 321-345.

Shafrir, U. (2001). *Meaning equivalence: Theory and applications*. Unpublished manuscript, University of Toronto, Toronto, Ontario, Canada.

Shafrir, U. (1999). Representational competence. In I. E. Sigel (Ed.), *The development of mental representation: Theory and applications* (pp. 371-389). Mahwah, NJ: Erlbaum.

Sigel, I. E. (1999). Approaches to representation as a psychological construct: A treatise in diversity. In I. E. Sigel (Ed.), *Development of mental representation: Theories and applications* (pp. 3-12). Mahwah, NJ: Erlbaum.

SRI, International (2002). *GLOBE classroom assessment tools.* Downloaded September 5, 2006, from http://globeassessment.sri.com/index.html.

St-Cyr, O. (2002). *An Ecological Approach to Mental Models: Towards Assessing Ecological Compatibility*. Unpublished MASc Thesis, University of Waterloo.

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.

Thagard, P. (1989). Explanatory Coherence. *Behavioral and Brain Sciences, 12,* pp. 435-67.

Toulmin, S. E. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

Tversky, A. and Kahneman, D. (1980). Causal Schemata in Judgments Under Uncertainty. In M. Fishbein (ed.), *Progress in Social Psychology, vol. 1,* Hillsdale, NJ, Erlbaum, 49-72.

Tversky, B. (1997). Spatial constructions. In N. Stein, P. Ornstein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp.181-232). Mahwah, NJ: Erlbaum.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Vera, A. H., & Simon, H. A. (1993). Situated Action: Reply to Reviewers. *Cognitive Science: A Multidisciplinary Journal* 17:1, 77-86 von Glasersfeld, E. (1987). Preliminaries to any theory of representation. In C. Janview (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 215-225). Hillsdale, NJ, Lawrence Erlbaum Associates.

Vygotsky, L. S. (1977). Denken und Sprechen [Thought and Language]. Munchen: Fischer Taschenbuch Verlag.

Welbank, M. (1990). An overview of knowledge acquisition methods. *Interacting with Computers, 2,* 83-91.

Sponsor
The National Science Foundation, Grant REC-0129331

Prime Grantee
SRI International. *Center for Technology in Learning*

Subgrantees
University of Maryland
University of California, Berkeley. *Berkeley Evaluation & Assessment Research (BEAR) Center* and *The Full Option Science System (FOSS)*
University of Michigan. *BioKIDs*