How Structural Assessment of Knowledge can be used for the identification of specific alternative conceptions and for assessing domain competence in physics

Harold Sharara

University of Ottawa.

Thesis submitted to

Dr. David Trumpower (Thesis Supervisor)

Dr. Christine Suurtamm and Dr. Louis Trudel (Thesis Committee)

in partial fulfillment of the requirements for EDU 7999.

ACKNOWLEDGEMENTS

I thank God for helping me to complete this thesis. I am also greatly indebted to my thesis advisor, Professor David Trumpower, and my thesis committee. Thanks to my family for all the support (emotional, spiritual, material and financial). Special thanks to my wife for all the encouragement.

Mark 9 v 23.

CONTENTS

How Structural Assessment of Knowledge can be used for the identification of specific alternative conceptions and for assessing domain competence in physics.

ABSTRACT

The purpose of this study is to investigate the viability of Structural Assessment of Knowledge (SAK) as a tool for identifying alternative conceptions and for predicting domain performance in Physics. The process begins by eliciting and then representing students' knowledge. One of these types of knowledge is conceptual knowledge, which is important for performing procedural tasks. This thesis employs a cognitively based theoretical framework to uncover students' knowledge, and then represents that knowledge for analytical purposes using SAK. SAK uses the Pathfinder algorithm to empirically derive the semantic networks of the students' and experts' cognitive structures, by asking them both to rate the relatedness of pairs of physics terms. Comparing students' and experts' knowledge structures provided some support for the structural assessment theory. In particular, supporting evidence that Pathfinder networks help in predicting student's problem solving capabilities was attained.

Thesis Supervisor	Dr. David Trumpower
Thesis Committee:	Dr. Chris Suurtaam
	Dr. Louis Trudel

I. INTRODUCTION

Seventeenth century English philosopher John Locke suggested that students come to school as “tabula rasa” (blank slates) to be “written upon” by teachers. While Locke was correct about a great number of things, this was not one of them. Students come to school with ideas that deal with the natural world that are highly resistant to change and strongly influence new learning (Carmichael et al., 1990; Pfundt & Duit, 1991). Often these ideas or conceptions that students carry with them in a particular domain are different from those of scientists and/or experts in the field (Costu & Ayas, 2005). It is these differences that are collectively known as alternative conceptions. Science education research reveals that, even with instruction, students at the post-secondary level consistently possess deep-seated alternative conceptions in various scientific disciplines (Driver & Easley, 1978; Gardner, 1986; Griffiths & Grant, 1985; Hackling & Garnet, 1986; Halhoun & Hestenes, 1985). The children’s book Fish is Fish by Leo Lionni (1970) illustrates this problem beautifully. Gilbert and Watts (1983) cite Lionni telling a story about a fish that is interested in learning about life on land. Unfortunately, the fish cannot explore any place beyond the confines of a small pond. He befriends a tadpole that eventually grows into a frog and moves out of the pond onto the land. The frog subsequently returns to the pond and reports what he has seen to the fish. The frog describes all kinds of things such as people, birds, and cows. The book’s illustrations depict the fish’s mental representations (or mental model) of each of those things described by the frog; each land creature has a fish-like body that is slightly adapted to accommodate the frog’s descriptions. People are imagined to be fish that walk on their tailfins, birds are thought of as fish with wings, and cows are believed to be fish with udders. This children’s story exhibits well both the creative license and dangers inherent in the fact that people construct new knowledge based on prior experiences and understandings. Research has shown that instead of remembering a host of accurate details, people tend to remember events by incorporating a few details within a pre-existing schema for the event (Silva et al., 2006; Scoboria et al., 2006). Alternative conceptions often result when new experiences are interpreted in light of prior experiences, and new understandings are grafted onto prior understandings. Memories in general are retrieved by first recalling the schema and then the associated details. If a concept does not fit a pre-existing schema and is not all that salient, it likely will be forgotten or even rejected.

The present research, to identify these alternative conceptions, will be based on several theoretical perspectives as described below. It is assumed that students use mental models or explanatory frameworks (general concepts used to interpret and explain the world) to organize knowledge and enhance its retrieval during problem solving tasks. When these mental models are fragmented and/or inadequate, understanding is inhibited. Thus, one approach to enhancing students' understanding of physics is to help them acquire an appropriate conceptual framework as they construct their understanding of physics. The ideas, skills, and abilities that students bring with them are referred to as prior knowledge (Jonassen & Gabrowski, 1993; Tallman & Phene, 2007). Prior knowledge greatly affects how students understand, remember, and ultimately learn new information (Halikari, Nevgi, & Lindblom-Ylanne, 2007). It has also been suggested that non-domain-specific factors, for example, religious and cultural beliefs (Matsumura, 1998), difficulties with the scientific meaning of common terms (Zaim-Idrissi, Desautels, & Larochelle, 1993), and inadequate textbooks (American Association for the

Advancement of Science, 2000; McComas, 1997), may be factors in students' persistent alternative conceptions. Structural Assessment of Knowledge (SAK) is proposed here as a technique that may help to identify alternative conceptions.

II. REVIEW OF CONTEMPORARY RESEARCH

A basic assumption in Cognitive Psychology is that knowledge requires an understanding of the interrelationships among concepts and that this organization of knowledge can best be analyzed with structural representations (Jee & Wiley 2007; Mahon & Caramazza, 2007). Recent theories of learning implicate the structure of one’s knowledge in recall, inferencing, comprehension, and problem solving (Baxter, Elder, & Glaser, 1996; Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004; Trumpower & Goldsmith, 2004). As a result, this review will take a close look at how knowledge is acquired, the problems associated with knowledge acquisition, and finally discuss some of the ways to represent and assess knowledge.

1. ACQUISITION OF KNOWLEDGE AND ITS ORGANIZATION:COGNITIVE MODELS OF LEARNING

Several researchers (Chi, 1993; Chi, Slotta, & deLeeuw, 1994; Vosniadou, 1994; Ferrari & Chi, 1998; Slotta & Chi, 1999) have proposed that learners use mental images (or explanatory frameworks) to "make sense" of many scientific phenomena. These deeply ingrained images act as facilitators or inhibitors of future knowledge retrieval and problem solving. For example, students often think of force as a kind of impetus imparted to a body or as an intensive property that a body possesses (similar to velocity). It is known that some students believe this force must be continuously applied to an object in order for the object to remain in motion and that heavier objects fall faster (Halloun &

Hestenes, 1985). They assume that this impetus can be used up. Such alternative conceptions typically show up in students' attempts to understand Newton's laws of force, resulting in the familiar assumption that when there is no force, there is no movement (because of the absence of impetus). By the same token, we can predict that phenomenon for which the learner's conceptions and the scientific conceptions are compatible should be relatively easy and straightforward to learn. Take, for example, the biological concept of *animal***.** Children's conceptions about the concept *animal* stem from representing it as a variant on the concept of *human* (Carey, 1985). But this conception does not involve an incompatibility because an adult (or the scientific) conception, *human*, belongs to the more super ordinate *animal* category. Therefore, children acquire a fairly sophisticated understanding of the concept of *animal* by age 10 years, as Carey's (1985) data, and more recently, Massey, Freyd, and Roth's (1992) data have shown.

2. PROBLEMS ASSOCIATED WITH THE ACQUISITION OF KNOWLEDGE AND ITS ORGANIZATION

Piaget (1954), Ausubel (1963 & 1968), and Novak (1988) propose a representational model in which domain-specific declarative knowledge is stored as a network in long-term memory. These network models assume that concepts are stored as nodes that are interconnected to form a vast associative network. The declarative knowledge or semantic features refer to lower-level knowledge that does not transfer as widely as conceptual knowledge, while mental models are analogies of events or situations (declarative knowledge), together with procedural rules to mentally manipulate the events or situations. Declarative knowledge involves simply knowing facts and ideas (*what*), whereas on the other hand, procedural knowledge involves knowing the way to do a task (*how*). Individuals use mental models to predict future events, answer comprehension questions, or solve problems. Initially, individuals construct mental models that include only declarative knowledge; however under appropriate conditions, learners put less emphasis on these semantic features and construct mental models of the situation, by encoding procedures, goals, and relationships (McNamara, Miller, & Bransford, 1991).

Knowledge structures are based on the premise that people organize information into patterns that reflect the relationships that exist between concepts and the features that define them (Johnson-Laird, 1983). In addition to the term *knowledge structures*, several other labels have been attached to the construct of knowledge organization, including *mental models*, *schemas*, and *conceptual frameworks* (Dorsey, Campbell, Foster, & Miles, 1999). Knowledge structures can be distinguished from declarative and procedural knowledge. They represent the organization of knowledge, whereas declarative and procedural knowledge reflects the amount of knowledge or facts or procedures learned. Measures of cognitive learning outcomes traditionally have involved achievement tests that assess declarative and/or procedural knowledge (Alliger et al., 1997; Goldsmith & Kraiger, 1997). However, current thinking in cognitive science suggests that, in addition to the amount of knowledge stored in memory, the organization of knowledge stored in memory is perhaps of equal or greater importance (Ruiz-Primo, Shavelson, & Wiley, 2005; Mahon & Caramazza, 2007). Well organized knowledge structures provide individuals with a responsive mechanism for retrieving information from long-term storage. Effective retrieval of information enables faster and more complete comprehension, enhanced capability to make inferences, prediction of future events, and determination of optimal actions that will influence current and future events in a desired way (Collins & Gentner, 1987). The present study extends this body of research by examining the relationship between participants' knowledge structures and the acquisition of problem solving skills. It is expected that the quality of participants' knowledge structures will have a positive correlation with problem solving transfer. As individuals develop expertise in a domain, their knowledge structures converge toward a true representation of that domain (cf. Acton, Johnson, & Goldsmith, 1994). Assuming that experts' organization and comprehension of domain knowledge are a close approximation of the true representation of that domain, then similarity to an established expert structure can be considered an indicator of skill development.

3. ASSESSMENT OF KNOWLEDGE ORGANIZATION AND ALTERNATIVE CONCEPTIONS

Ausubel (1968) and DiCerbo (2007) have noted, “the most important single factor influencing learning is what the learner already knows.” The previously held conceptions are variously referred to as, “preconceptions,” “alternative conceptions” and sometimes “misconceptions”, but the central idea is of conceptions that (i) are strongly held; (ii) differ from expert conceptions; (iii) affect in a fundamental sense how students understand natural phenomenon and scientific explanations; and (iv) must be overcome, avoided, or eliminated for students to achieve expert understanding (Hammer, 1996). Note that not all conceptions that differ between student and expert are included in this description, only ones that pose an obstacle to learning. The first step in dispelling such alternative conceptions is to identify them and to recognize their sources. To identify alternative conceptions, researchers and instructors resort to various methods including the following:

- using open-ended questions to assess what students know about a topic

- listening to, and observing students' answers
- using direct questioning to discover the students' reasoning process
- analyzing responses to multiple choice questions
- interviewing students and/or instructors.

Instructor interviews have been employed in the domains of statistics (Khazanov, 2006) and elementary students' physics knowledge (Pine, Messer, & St. John, 2001) to successfully develop lists of common alternative conceptions. Boo and Watson (2001) conducted interviews with students each year for three years to monitor their progression in eliminating fundamental alternative conceptions about chemical reactions. Others have used open-ended questions to elicit alternative conceptions from students (Costu & Ayas, 2005). These approaches allow in-depth information to be gathered from students, but are time consuming and not easily used with large numbers of students. In order to identify common alternative conceptions about physics, Halloun and Hestenes (1985) constructed a multiple choice test on which different answers related to different possible ways of theorizing about the problems. By administering this to over 400 students, they began to understand that various ways of thinking were a common occurrence. Following the identification of common alternative conceptions, researchers often attempt to create an assessment to measure the extent to which students hold these alternative conceptions. The Force Concept Inventory (FCI) (Hestenes, Wells, & Swackhamer, 1992) and Force and Motion Conceptual Evaluation (Thornton & Sokoloff, 1998) both have been used in physics classes to measure whether students understand physics concepts that underlie the calculations necessary for solving physics problems. One of the weaknesses associated with this technique, however, is that multiple choice tests are susceptible to guessing. The results can be misleading, for example if subjects guess correctly, giving the illusion that they in fact have a correct understanding of the concepts underlying physics calculations. Another concern with the techniques is whether they are easily generalizable to other domains besides physics. This issue makes a good argument for using SAK because of its relative ease to be applied in different domains.

Students with alternative conceptions construct different mental models by which they explain concepts and processes happening around them. As shown in the Carey (1985) example above, they do so already as young children and before receiving formal instruction on those concepts. These mental models usually contain alternative conceptions that are far from scientifically correct (Diakidoy, Vosniadou, & Hawks, 1997). Many students have difficulty in changing their alternative conceptions because these false concepts may be deeply ingrained in the mental map of students or they are committed to their existing knowledge (Reiner, Slotta, Chi, & Resnick, 2000).

The conceptual change approaches of Ausubel (1963, 1968) and later of Novak (1985) marked the beginning of the most active period of research on alternative conceptions. A conceptual change is a change in the way individuals experience the world; the result of meaningful learning. Ausubel distinguished between *meaningful* and *rote* learning. The underlying basis of the theory is that meaningful (as opposed to rote) human learning occurs when new knowledge is consciously and purposively linked to an existing framework of prior knowledge in a non-arbitrary, substantive fashion. In rote (or memorized) learning, new concepts are added to the learner's framework in an arbitrary and verbatim way, producing a weak and unstable structure that quickly degenerates. Ausubel rejected the verbatim repetition of definitions and the reproduction of procedures and specified the conditions for meaningful learning: appropriate materials, the motivation of the student to relate old and new ideas, and preconceptions, which encourage the student to revisit his preconceptions. Ausubel saw cognitive structure as hierarchically organized. He used the terms *subsumption,* the use of general concepts to acquire and organize new concepts; *progressive differentiation,* in which new links are formed among related concepts; and *integrative reconciliation,* in which a learner relates previously held concepts to resolve conflicts in meanings. This proposition of Ausubel makes it more important to examine what a student already knows.

Clement's studies provide a classic example of research on alternative conceptions in the early phase. Clement (1987b) suggested that alternative conceptions might provide a good starting point for instruction and that a theoretically consistent model can be approached by means of bridging analogies. Clement conducted a tutorial in which he tried to bridge the gap between examples already known to students, what he called *anchoring examples*, to illustrate certain laws or principles, and the *target examples* which the students did not perceive but that imparted the same idea. The *bridging examples* were used as supports in conceptually bringing the two analogous examples closer together. To explain further, consider the common alternative conception, contrary to Newton's third law of motion, that a table does not exert a force on a book lying on it even though students understand that the book exerts a force on the table. The anchoring example which could activate the correct schema would be to ask students to use their hand to push down on a spring. When students would exert a force downward on the spring, they could also feel the spring exerting a force upward on their hand. Clement (1987b) found that almost all students understand the anchoring example that a spring can push back on their hand, but many were still unconvinced that this is a valid analogy to the target example which is the book on the table. This is because the anchoring example is too far removed from the target example and so intermediate examples which serve as bridges must be given. These might be the book on the spring, then the book resting on a piece of foam, and then the book resting on a flexible board. In this progression, students gradually move away from the anchor and towards the target.

Research in science education shows consistently that students with alternative conceptions are relatively unaware of conflicts between their knowledge and the new information they are receiving (Guzzetti, Snyder, Glass, & Gamas, 1993). When the teacher asks students with alternative conceptions to elaborate a concept, they integrate their background knowledge as do students without alternative conceptions (Hannon & Daneman, 2001). But the content of knowledge-based answers in the domain reflects the alternative conceptions and students are only able to give fewer valid explanations. Such a lack of explanations can undermine students' success in other tasks such as problem-solving in science or understanding expository texts (Ohlsson, 2002). Stavy and Tirosh (1996; 2000) examined common thinking patterns of students having alternative conceptions in different areas. They found that by creating logical environmental relations, students try to understand the given situation to a certain extent without having a real understanding of the concept.

A question arises here: How do different students understand knowledge and knowing? The nature of students' understanding is best expressed in the following quotation: Epistemology is always *personal*. Research on students' personal epistemologies (Hofer & Pintrich, 2002) has conceptualized alternative conceptions as stable beliefs or stages of development. Hammer and Elby (2002) proposed that personal epistemologies are comprised of manifold epistemological resources. Some students believe that (1) understanding physics means learning formulas and facts; (2) knowledge of physics is loosely related with what happens in the real world; and (3) learning physics means memorizing content of the physics textbook. In contrast, other students may believe that (1) understanding physics means developing a sense of the basic principles underlying it; (2) the principles of physics explain the phenomena of the physical world; and (3) the process of learning physics leads to modifying one's own understanding. Although students activate and use their prior knowledge during the learning process, due to the alternative conceptions they fail to incorporate new information into their prior knowledge. Research on students' alternative conceptions also shows that they are resistant to change (Hameed, Hackling, & Garnett, 1993; Baser & Geban, 2007) which affects students' comprehension of new knowledge (Kendeou & van den Broek, 2005; Diakidoy & Kendeou, 2001). This results in the compartmentalization of old from new knowledge as demonstrated earlier by the children's book example, Lionni (1970). If new knowledge does not fit into a pre-existing schema it will likely be rejected or compartmentalized from the old. Such compartmentalization is a crucial educational problem. How does a teacher nurture incorporation rather than compartmentalization? An important prerequisite would be that the teacher is aware of inconsistencies between students' prior knowledge and new information they are learning (Kendeou & van den Broek, 2005). Understanding alternative conceptions and the effect of specific aids on developing students' knowledge structures is very important. Teachers must address alternative conceptions if they want to alter them. There has been considerable interest in attempts to design instructions to facilitate cognitive structure change, and to address alternative conceptions (Ausubel, 1963, 1968; Novak, 1985; Treagust & Duit, 2008). Information about students' alternative conceptions in a class can assist the teacher in preparing effective lesson plans (Atkinson, 2004).

Clearly, different methods of identifying alternative conceptions have different strengths and weaknesses. The methods above largely center on alternative conceptions related to specific topics. However, in reality students must learn not just individual concepts but how many different concepts fit together (Ausubel, 1968). Just as a student may understand how to do a calculation without understanding why, a student may understand one component of a domain without understanding how it fits in to the domain (Borgen & Manu, 2002). The following section presents a discussion of how to assess knowledge of conceptual relationships, or conceptual structure.

4. ASSESSING CONCEPTUAL STRUCTURES

The central idea behind a cognitive assessment is the link between an individual's competence in a domain of learning and how to represent the subject matter of the domain. Presumably, an important characteristic of knowledge in any field of study is its structure or organization. The concepts, ideas, terms, and rules that make up a domain are interrelated in ways that give rise to their meaning and the meaningfulness of the domain (Goldsmith & Kraiger, 1997). Methods of inferring the way in which people organize domain-specific information are very diverse (Adelson, 1981; Egan & Schwartz, 1979; Murphy & Wright, 1984; Shavelson, Shoenfeld & Herrmann, 1982). They are all based on representational models of memory. They all maintain that conceptual structures are symbolic internal representations of external reality stored in long-term memory. They are based on strong assumptions that conceptual structures are relatively stable (once learned) and meaningful (have semantic properties), and can be inferred from an individual's overt behavior. These methods can be categorized as verbal reports, clustering methodologies, and scaling techniques (Koubek & Mountjoy, 1991). Each method involves three stages: eliciting the knowledge that has been acquired, generating a representation of the conceptual structure, and quantifying the quality of conceptual organization.

Concepts and their relationships can be elicited by interviews, essays, concept maps, ordered lists, and pair-wise similarity ratings such as described below in Pathfinder Networks. Educators and cognitive psychologists have developed several methods of then converting elicited relationships into representations of the way in which knowledge is structured (Olson & Biolsi, 1991). Verbal data derived from interviews or essays can be analyzed in order to identify parts of speech and propositions or syntactic relations, which are transformed into a network of labeled interconnected nodes such as concept maps (d'Apollonia, De Simone, Dedic, Rosenfield, & Glashan, 1993; Frederiksen & Breaueux, 1990; Mosenthal & Kirsch, 1992). For example, Novak and Musonada (1991) interviewed students on their understanding of chemistry, converted their protocols into concept maps, and subsequently assessed the maps. Clustering techniques, such as K-means Cluster Analysis can be used to transform ordered lists into network representations. And finally, similarity ratings, and other such quantitative proximity data can be transformed via scaling techniques such as Pathfinder (Schaneveldt, 1990) or by multidimensional scaling, into general weighted networks. Thus, there are many different techniques available for measuring knowledge structures.

Structural Assessment (SA) is considered to be the modal technique for measuring knowledge structures (Kraiger et al., 1993). The present research demonstrates how SA can be used to quantify knowledge structures. In addition, the study illustrates how SA yields representations of organized knowledge that reflect comprehension of domain concepts and can be related to procedural tasks. Although tests of declarative knowledge traditionally have been used to assess learning after training, empirical evidence over the past 20 years suggests that measures of knowledge structures provide more specific information about an individual's understanding. In a training context, knowledge structures reflect the degree to which trainees have organized and comprehended the content of training. Indeed, research has indicated that measures of knowledge structures using SA can differentiate between experts and novices (e.g., Schvaneveldt et al., 1985), predict classroom learning and achievement (e.g., Goldsmith, Johnson, & Acton, 1991), reflect training manipulations designed to influence learning (Kraiger, Salas, & Cannon-Bowers, 1995), and predict transfer performance on tactical decision-making tasks (Kraiger et al., 1995).

Concept Maps

To support the argument for one particular SA technique called Structural Assessment of Knowledge (SAK), which is the main focus of this study, it is necessary to initially review another more commonly known SA technique, concept maps, in this section. Here, we will uncover some of the shortfalls of concept–mapping that are later addressed by SAK. This is intended to build a case for SAK as an alternative assessment technique.

Some educators propose assessing structural knowledge with the use of concept mapping, by which students denote structural relations among concepts through hand-drawn or computer-drawn maps (All, Huycke, & Fisher, 2003; Williams, 2004). Concept maps are visual representations of how concepts are related to each other. Each concept is a node, connected to other related concepts. The maps theoretically represent the knowledge structure of the person who is drawing them. There have been a variety of methods created to score concept maps, including holistic scoring (based on overall understanding of the concepts represented by the map), density scoring (based on the number of nodes and links), and validity scoring (based on “correct” links), all with or without comparison to a master map (McClure, Sonek, & Suen, 1999). Different researchers have also assigned different point values or weights to different map elements. There have been concerns expressed about the reliability of these scoring methods (Ruiz-Primo and Shavelson, 1996), although others have reported acceptable inter-rater agreement (e.g., Rice, Ryan, and Samson, 1998). However, even if reliable, it is still time-consuming to score concept maps (McClure, Sonak, and Suen, 1999).

Most of the scoring mechanisms described above were designed in an effort to get an overall assessment score, not indicate areas of alternative conceptions, so they tend to produce a single score for a map. This is useful for summative assessment purposes, but does not provide enough information to identify alternative conceptions. Although the single score is useful in differentiating between students with more correct concepts from those with fewer correct concepts, it says little about the differences between students with identical scores, for example. Do students with the same scores, have the same conceptual relationships? Not necessarily! The limited information provided by the single score presents a problem for comparing maps with the same score. In fact, much of the literature on concept mapping indicates its primary usefulness as an instructional strategy. When used for evaluation, scoring of concept maps can be subjective and time consuming.

(a) The SAK technique

One technique has emerged that may allow for both the investigation of large numbers of students and the location of specific alternative conceptions - the use of numerical judgments of relatedness or closeness among a set of concepts (Kraiger, Ford, & Salas, 1993). These ratings can then be converted into concept map-like networks with unlabeled links, and their structures assessed using procedures in the Pathfinder network program (Schvaneveldt, 1990). Collectively, this technique is known as SAK.

The Pathfinder network scaling algorithm generates a network representation of the structure of the relatedness ratings submitted for analysis. This representation, referred to as a Pathfinder network or PFnet, contains concepts represented by nodes, and links between nodes to demonstrate relatively strongly perceived relationships between concepts (Trumpower, Sharara, & Goldsmith, 2010). Figure 1 below is an example of a PFnet produced from relatedness ratings of experimental design concepts.

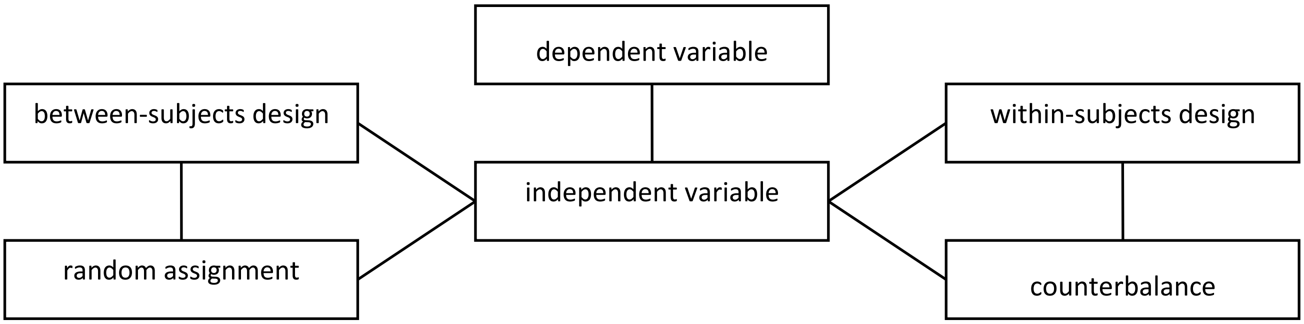

Figure 1. Pathfinder network with experimental design concepts.

The relatedness ratings that are submitted to Pathfinder are treated as proximities as in Figure 2 below. The algorithm then proceeds to identify the shortest indirect (or direct) path between each pair of concepts using the proximity ratings.

Relatedness Rating	Pathfinder Proximity Rating
7	1
6	2
5	3
4	4
3	5
2	6
1	7

Figure 2. Relatedness and proximity ratings.

A PFnet will only include a direct link between two concepts if the shortest indirect path between them is greater than the direct path. For example, if the proximity between “A-B” is 2; “B-C” is 3; and “A-C” is 6, then Pathfinder will not include a direct link between “A” and “C” because the most direct path (shortest indirect) between them is through “B” (2+3=5). Pathfinder computes the links by computing all paths between nodes and including a link between nodes only if it represents the shortest path. The resulting PFnet can then be compared to an expert/ instructor’s PFnet and a measure of similarity obtained based on the number of common links. The basic steps required to implement SAK are outlined in the following sections.

i. Elicitation

The elicitation process must capture structural knowledge, i.e. the relationships among important subject matter concepts. Various techniques have been used to obtain concept relationships, including word associations (Johnson, 1964), card sorting (Shavelson & Stanton, 1975) and direct numerical rating of pair wise relatedness (Fenker, 1975; Goldsmith et al. 1991). Pathfinder makes use of this numerical rating procedure by asking subjects to make their judgments of relationships between concepts on a relatedness scale. For example, a 7 point scale might be anchored by 7 = highly related, 1 = not at all related. Some advantages of this elicitation method are that (a) the validity of direct pair wise relatedness ratings has been demonstrated (Cooke et al, 1986), (b) the method is easy to explain to learners, (c) the judgment task is simple and (d) the resulting data can be directly represented using the methods described next.

ii. Representation

A physical portrayal of the structure of concepts is required. An array of relatedness ratings can be transformed by a scaling algorithm to derive a meaningful representation. The assumption in a scaling approach is that the inherent structure is derivable from the perceived similarities of a set of items. Several scaling methods have been used to derive representations, including multi-dimensional scaling (MDS) used by Kruskal (1964). The pathfinder scaling algorithm has proved to be another useful method for representing knowledge (Schvaneveldt, Durso, & Dearholt, 1989), and is usually preferred over the other methods. For example, MDS or cluster analysis also creates visual representations of the structural framework underlying a set of concepts (Fenker, 1975). However, MDS is limited to symmetrical data, or concepts that are linearly related, and cluster analysis is limited to class-inclusion structures or concepts that are related by position. Pathfinder transforms the matrix of similarity ratings into a structural network where the concepts are depicted as nodes and relatedness between concepts is depicted by how closely they are linked. Gonzalvo, Canas, and Bajo (1994) combined the Pathfinder analysis with multidimensional scaling and found that the multidimensional scaling represented the global properties of networks and the Pathfinder networks captured the local conceptual relations. This finding will become particularly important for this study later in Section III below. Pathfinder's other advantage over concept maps, but not MDS and cluster analysis, is that it does not impose a hierarchical solution, therefore allowing greater freedom in reflecting an individual's inherent cognitive structure.

iii. Evaluation

In a training context, evaluation may be used to assess pre-test post-test change or find the level of competence indicated by a particular cognitive structure. The evaluation may take the form of a qualitative assessment of the derived structure (Hamrick, Harty, & Ault, 1987; P.E. Johnson, Cox & Curran, 1970) or a quantitative assessment (Preece, 1976; Shavelson, 1972). The latter technique is usually preferred because it allows criterion referenced evaluation and provides specific feedback to learners. This aspect becomes especially important in formative assessments. In a quantitative analysis, one evaluation called Pathfinder Similarity (PFSim) compares two network representations (Goldsmith & Davenport, 1990). PFSim measures the configural similarity between two network structures containing identical concepts. Competence in an individual's cognitive structure is assumed if the structure is similar to the referent or target structure. In any two networks, PFSim varies from zero to one. If the networks are complementary (the presence of a link in one implies the absence in the other), PFsim equals zero. If the networks are identical, i.e. for each concept, all concepts linked to that concept were identical, PFSim equals one. In addition, ratings can be combined to produce “average” networks across individuals. The Pathfinder technique has been used to evaluate changes in conceptualization before and after instruction in evolution (d’Apollonia, Charles, & Boyd, 2004), statistics (Geske, 2001), and police training (Braverman, 1997). Additionally, Pathfinder scores have been shown to predict achievement (Chen, 1997; Goldsmith, Johnson, & Acton, 1991). Johnson, Goldsmith, and Teague (1994) suggest the predictive power of Pathfinder comes from its consideration of all of the possible links with a concept and its ultimate dichotomization into directly linked and indirectly linked concept pairs.

As just indicated, SAK requires that individuals’ knowledge structures be compared to that of an expert (or referent). The referent structure serves as a standard against which individual cognitive structures are evaluated. This structure is acquired by obtaining experts’ judgments of the relatedness of all possible pairs of the concepts to be evaluated. By applying the scaling technique to these judgments, the referent structure is then derived. Acton, Johnson, and Goldsmith (1994) showed that averaged ratings from multiple experts tend to have better predictive validity than representations obtained from single experts. They compared an average expert referent structure with that of individual expert referent structures in predicting students' exam performance and discriminating between various levels of expertise (i.e., novice and advanced students). They indicated that, despite substantial variability between individual expert structures, the average expert referent structure provided predictive validities that were stronger than the predictive validities for the individual expert referent structures and greater discrimination between novice and advanced students. They also indicated that the average referent structure provided more stable predictive validities than the validities of the individual expert referent structures. Furthermore, the individual expert structure that provided the strongest validities was not meaningfully superior to the average expert referent structure in predicting exam scores. They concluded that "experts do not have to agree that highly in an absolute sense to allow for averaging" (Acton et al., 1994, p. 310). Additionally, they advocated averaging individual expert structures to form a referent structure rather than searching for the ideal individual expert structure.

(b) SAK and Validity

It was previously noted that Pathfinder has been used for representing knowledge in a number of domains, including basic research in memory (Cooke, 1992; Cooke, Durso, & Schvaneveldt, 1986), learning (Gomez & Schvaneveldt, 1994), problem solving (Durso, Rea, & Dayton, 1994), representation of belief in psychosomatic illness (Gomez, Schvaneveldt, & Staudenmayer, 1996), representation of the cognitive structures underlying expertise (Cooke & Schvaneveldt, 1988), and of particular relevance to the present study, structural assessment of knowledge growth for academic subject matter (Acton, Johnson, & Goldsmith, 1994; Goldsmith, Johnson, & Acton, 1991; Gonzalvo, Canas, & Bajo, 1994; Housner, Gomez, & Griffey, 1993a, 1993b; Johnson, Goldsmith, & Teague, 1994). For example, Goldsmith et al. (1991) assessed the similarity between instructor and student PFnets in the context of a university course on statistics and experimental design. The degree of similarity between student and referent Pathfinder networks was a better predictor of performance than were correlations between untransformed relatedness ratings and exam scores, or measures based on the use of multidimensional scaling. Given previous research suggesting that multidimensional scaling captures global properties of the relations among concepts whereas Pathfinder captures the local properties (Cooke et al., 1986), these findings suggested that reducing the untransformed relatedness ratings to only the most salient relations among concepts, as is done in Pathfinder analysis, might be capitalizing on the configuration of domain knowledge. Goldsmith et al. (1991) argued that if knowledge is indeed based on sensitivity to the relationships among concepts, then a method that captures and represents the configural character of these relationships should be particularly useful. In short, these studies suggest that a structural-based approach can provide a useful means for representing and assessing the growth of domain knowledge. Relatedness ratings represent a more indirect, but possibly objective, approach to discovering the structure of knowledge than do standard exams. Students are required to make a judgment about every pair of concepts; thus, they are tested over a wider range of comparisons than would be possible on an essay or multiple-choice exam. Since there is no context provided other than that which arises from the relationship between the two concepts, relatedness ratings are not restricted to nor biased by the context implied by the test question. Relatedness judgments also rely greatly on deeper understanding and thus are less susceptible to the problems of rote memorization. Another problem with standard exams has to do with devising simple and objective grading systems. This is especially true in the case where student knowledge is assessed relative to other students rather than directly in relation to the course instructor. Thus, some traditional exams may test the factual knowledge of one student relative to another but give little insight into how students organize knowledge in relation to the instructor or some other referent

(Goldsmith et al., 1991). On this view, the use of relatedness ratings in combination with techniques for determining the most salient relationships among the concepts may provide a means for better reflecting a system of knowledge than would the more traditional method of eliciting answers to specific questions on an exam.

Evidence for the validity of PFnets as a measure of structural knowledge is based on the assumption that there is some ideal knowledge organization that best reflects the structure of the domain and that cognitive structures become more like this ideal structure as experience in the domain increases (Acton, Johnson, & Goldsmith, 1994). Two frequently used approaches to address validity are demonstrations of increase in the similarity of a pathfinder network to a referent network after instruction, and demonstrations that the similarity between students' and expert's pathfinder networks are positively related to other measures of classroom achievement, such as course grades (Gomez, Hadfield, & Housner, 1996), or examination scores (Azzarello, 2007; d'Appolonia, Charles, & Boyd, 2004; Goldsmith, Johnson, & Acton, 1991). Several studies have shown that there is a positive change in students' PFnets before and after instruction. For example, research demonstrated that students' PFnets become increasingly similar to an instructor's PFnet from pre- to post-intervention in the field of pulmonary physiology (McGaghie *at el.*, 1996; McGaghie, McCrimmon, Mitchell, Thompson, & Ravitch, 2000). In another study, Azzarello (2007) investigated 102 students of health nursing during an 8-week course focusing on concepts related to nursing and public health strategies for populations, by using pathfinder techniques. He reported that the mean similarity of the students' PFnets to the referent network increased significantly from 0.19 before the course to 0.24 after the course completion. Findings were similar in studies of computer programming concepts (Trumpower, 2004) and computer networking (Dicerbo, 2007). Several studies have also shown a positive relationship between the similarity of students' and instructor's networks and students' test scores. Goldsmith, Johnson, and Acton (1991) found that the similarity between each student's PFnet and the instructor's PFnet was highly predictive of the examination scores over the course of the semester. In another study, Azzarello (2007) found that students with structural knowledge that was coherent and similar to the instructor's performed better in the class. The students' structural knowledge also differentiated between high-performing and low performing students. For example, in a study of 53 undergraduate students enrolled in a teaching methodology course, researchers examined the similarity between students' and instructor's PFnets created from ratings among 27 domain concepts. The correlation between the similarity of students' and instructor's PFnets and test averages was significant (r = 0.52) (Gomez *et al.,* 1996). The study conducted by Azzarello (2007) examined the similarity and correlation between PFnets for 102 community health nursing students and an instructor's network for a course. There was no significant correlation between the mean examination scores and pre-course similarity. The mean examination score was positively correlated with post-course similarity (r = 0.29).

The stability of experts' pathfinder networks may vary depending on the context of the data collection task (Gammack, 1990). The meaning of a concept depends to some degree on the particular context of use, and this relationship may vary across contexts, and relationships strong in one set of circumstances may not apply in another. For example, a liger is considerably more central if related to a zoo while less central if related to jungle animals. In a series of studies examining the stability of experts' PFnets of train locomotives, Gammack (1990, p. 226) argues, "across different elicitation conditions, no unique conceptual structure emerged when it might have been expected. Methodological and psychological reasons were considered as explanation for this instability, and with hindsight, it seems sensible to expect structural variation for a variety of reasons". The actual circumstances and any pre-existing schemas may emerge uncompromised, which influence and may modify the applicability of a given relationship. The instability of knowledge representation is neither caused by, nor unique to Pathfinder, which provides a tool for investigating the sources of instability in conceptual representation, such as individual theories and biases. Gammack further argues that despite some variations, Pathfinder is particularly useful in contexts where experts' domain conceptions are concise and remain relatively constant.

Housner et al. (1993b) found similar results in a study involving prospective teachers enrolled in pedagogy courses. Knowledge of key pedagogical concepts was organized more consistently and was more similar to the instructor's organization after completion of the course than it was in the beginning of the course. Furthermore, Housner et al. (1993a) demonstrated that the prospective teachers' knowledge showed the same growth trend in relationship to the referent structures of other teacher experts, and to a composite network based on these teachers, thus adding significantly to the generalizability and content validity of this research (see Acton et al., 1994, for similar results in the domain of computer programming, as well as a more thorough treatment of the issue of appropriate referent structures). Finally, Gonzalvo et al. (1994) used the configuration of Pathfinder networks to predict the acquisition of conceptual knowledge in the domain of history of psychology. Gonzalvo et al. conducted a fine-grained analysis of the relationship of the concepts in Pathfinder networks to students' abilities to define these concepts. Each concept in the instructor and student networks was assessed structurally in terms of the concepts to which it was directly linked. Well-structured concepts in the student networks were defined as those sharing greater degrees of similarity (in terms of direct links to the same set of concepts) with those links found in instructor networks. Ill-structured concepts were defined as those sharing less student-instructor link similarity. Gonzalvo et al. found positive correlations between the goodness of students' definitions of concepts and structural similarity of their concepts to the concepts in the instructor networks, as well as an increase in the number of well-structured concepts at the end as compared with the beginning of the course.

Research also provides validation for the specificity of information that links in PFNets give. To illustrate this point, consider the study conducted by Dayton, Durso, and Shepard (1990) who showed participants the following riddle: “A man walks into a bar and asks for a glass of water. The bartender pulls a shotgun on the man. The man says, “thank you” and walks out. What missing piece of information would cause the puzzle to make sense?” In order to solve the riddle, a participant would need to realize that the glass of water and the bar tender’s shotgun, were both remedies for the man’s hiccups. By way of a relatedness rating task, participants’ structural knowledge of the riddle was assessed with SAK. Fourteen concepts deemed appropriate to the riddle were obtained for purposes of generating PFnets for all participants. However, instead of deriving the usual global (or overall) measures such as PFSim, the authors analyzed specific links between concepts that were deemed crucial for solving the riddle. Comparison of the

PFnets of those who solved the riddle to those of the non-solvers revealed that solvers contained a link between *remedy* and *glass of water*, a link between *remedy* and *surprise*, and a link between *surprise* and *shotgun*. Non-solvers did not have these crucial links. Therefore, whether this specific subset of links in the PFnets was present or not, predicted the participants who solved the riddle. Similarly, Trumpower, Sharara, and Goldsmith (2010) also showed that a more fine-grained evaluation of PFnets derived from SAK can be used to identify learners' specific strengths and weaknesses. The presence of particular links in students' PFnets was associated with their performance on related types of computer-programming problems. In particular, evidence for the convergent and divergent validity of two subsets of links in discerning performance on particular types of computer programming problems was obtained.

III. GOALS AND THEORETICAL ORIENTATION

The premise of this study is Structural Knowledge Theory which generally states: to be knowledgeable within a domain requires individuals understanding and/or acquiring the structure of the domain's content. The study proposes to explore the efficacy of the SAK technique in exposing alternative conceptions and the development of problem solving performance in the context of physics. It seeks to determine if the links in students' PFnets, derived from their concept relatedness ratings, predict the *specific* kinds of errors that they make when solving any given problems. The quality of the students' PFnets is a function of a quantitatively derived measure. The overall similarity, PFSim, of a student's Pathfinder network to a referent network, as we read earlier, has been used and validated in previous studies and will also be used here. It is hypothesized that this overall similarity between the student's Pathfinder networks and the referent PFnet would predict their overall performance on the physics problems. But, although previous studies have shown this overall similarity to be a valid measure of domain knowledge (e.g., Goldsmith, et al., 1991), it is not primarily a diagnostic tool.

The importance of this study in analyzing and comparing individual links is realized by observing that while two networks may have the same overall similarity coefficient with respect to the referent network, the specific links they each share with that referent might be different. For example, figures 3 and 4 below show two hypothetical PFnets that have the same global or overall PFsim when evaluated against Figure 1 above. If we assume that figure 2 is a referent network, then it becomes clear that a global PFsim is not particularly diagnostic at very specific levels of analysis.

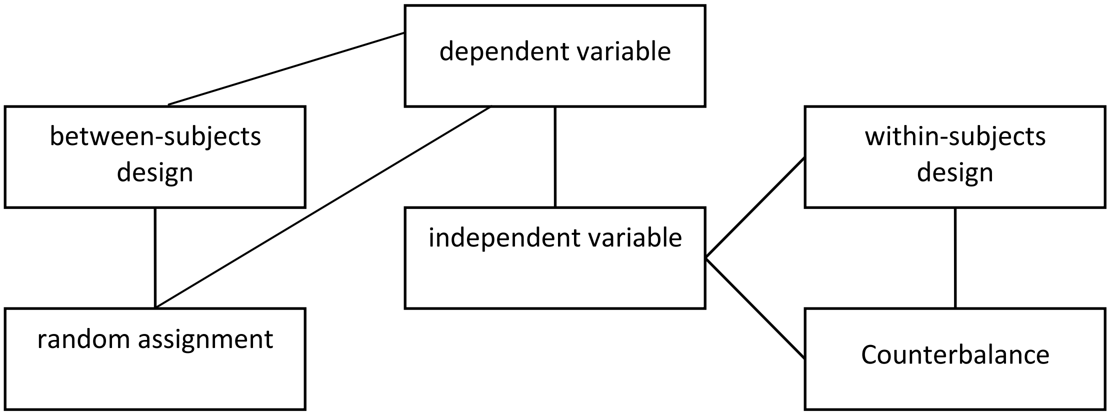

Figure 3. Pathfinder network of hypothetical student with stronger knowledge of within- than between-subjects designs.

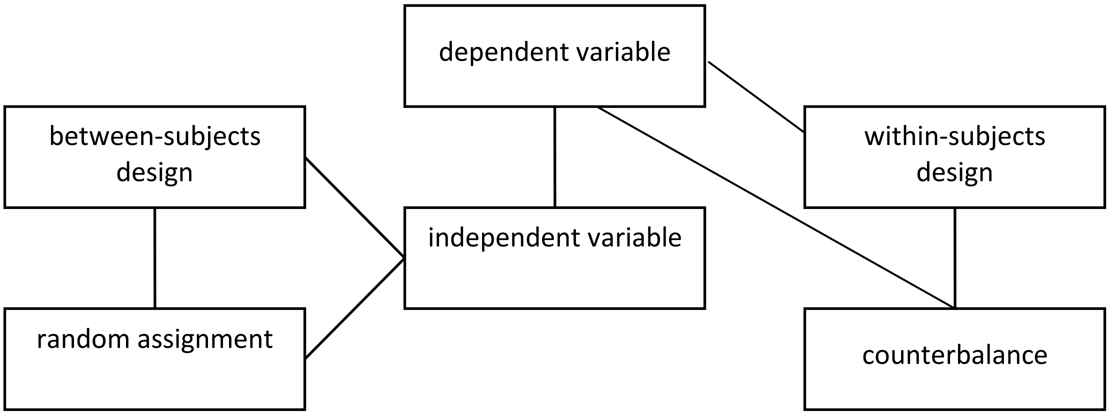

Figure 4. Pathfinder network of hypothetical student with stronger knowledge of between- than within-subjects designs.

Most importantly we discover that while both PFnets have a PFsim of .78 when evaluated against the referent network, they do not share the same links. The individuals' perceived understanding of experimental design conceptual relationships are not congruent. Since the two individuals have different relational linkages, it is reasonable to assume they would confront unique challenges when dealing with 'within-subjects' as opposed to 'between-subjects' designs. Hence, the additional need to analyze each student's cognitive structure through specific subsets of links and also one link at a time, and comparing this substructure to the referent network. This gave rise to the present study in which it is further hypothesized that similarity between subportions of the student's and referent PFnets will predict their performance on specific types of physics problems.

Overview of Present Study

The present study involved a secondary analysis of data from the University of New Mexico. The UNM research was interested in examining the effects of various training conditions on the development of both structural knowledge and problem solving performance. In general, students studied six solved physics problems, performed a concept relatedness rating task and then solved old, near and far transfer problems. Students' ratings were transformed using the Pathfinder scaling algorithm to generate a network representation of each participant's structural knowledge. These representations were then analyzed to determine differences in the organization of expert and novice knowledge. It was hypothesized that schemas would be identified in experts but not novices that are organized around general physics principles.

In this study, the ratings data is used for calculating PFsims at three specific levels of analysis, namely the overall, domain and equation levels. These PFsims are then correlated with problem solving performance at these same three levels to verify if performance could be predicted from PFsim, creating a unique difference from the original use of the data. This type of data analysis was the primary reason for this investigation and was not conducted in the original experiments. The results attained from this study were therefore independent of, and uninfluenced by, those attained in the original study. The following hypotheses guided the secondary analyses of this study:

Hypothesis 1: Students' overall PFsims will be positively related to their overall problem solving performance.

Hypothesis 2: Far Transfer problems require higher levels of conceptual understanding; consequently PFsims will have stronger positive correlation with far transfer than old or near transfer problem solving performance.

Hypothesis 3: PFsims based on specific subsets of concepts (e.g., those within a specific domain and within a specific equation) will be positively related to participants' performance on specific areas of problem solving (e.g., within a specific domain or use of a specific equation).

IV. POTENTIAL BENEFITS

SAK offers an alternative tool to identifying and correcting alternative conceptions in physics. Since SAK is not domain specific, one can appreciate the positive externalities that this study could impart to the fields of training and education. A diagnostic assessment tool is normally used as a pre-assessment tool at the beginning of a learning period. However, if SAK is considered as a tool for identifying alternative conceptions at any point within that learning period, then it can potentially be used for other purposes too. For example, if it proves to be a reliable diagnostic tool, then structural assessment can be useful for formative assessments. A formative assessment tool must be able to (a) identify students' precise knowledge gaps and/or misunderstandings, and to (b) provide feedback that can be used to fill the gaps and remediate the misunderstandings (Earl, 2003; McManus, 2006; Trumpower, Sharara, & Goldsmith, 2010). One of the things this study explores is whether or not SAK satisfies the first condition of a formative assessment tool – the ability to identify specific areas of strength and weakness. Therefore, if this study can prove the viability and usefulness of the SAK technique, it can help in the improvement of learning in other domains besides physics due to its easy adaptability.

V. RESEARCH DESIGN & METHODOLOGY

The original data was collected in 2003 by D.L. Trumpower as part of his doctoral dissertation. The main purpose of the study was to see the effect of various training manipulations on learning in physics. This was achieved through manipulating the type and number of word problems presented to students for studying, and the order in which they were presented. Presumably because the so called STEM subjects (science, technology, engineering and mathematics) have often “been considered among the most difficult subjects”, the researcher explored the underlying reasons for the development of expertise in the sciences. Initially, the study examined what characterizes expertise; “that is, how do experts differ from novices?” The findings suggested that experts’ knowledge is organized differently from that of novices. Consequently, the research went on to consider the differences between experts and novices in terms of knowledge organization. The next step was to examine various theories of schema acquisition; and finally, several training manipulations derived from those theories were tested. “Both problem solving transfer and a measure of knowledge organization that was independent of problem solving performance were implemented.”

In this study, I have used the data from both these measures to test the predictiveness of SAK on problem solving performance. In addition, I tested this predictiveness at three specific levels of analysis. SAK (overall) and problem solving performance were used as different outcome measures of learning in the original study. Thus, the relationship between problem solving performance and SAK was not determined, nor was structural knowledge or problem solving performance assessed at more specific levels as done here.

Participants

The participants were pooled from the University of New Mexico. Forty-five male and female students from the Psychology Department and three instructors from the Physics Department participated in the experiment. Student participation was voluntary and limited to those students who had not taken more than a single college-level physics class. Although mathematical proficiency varied among students, none had taken more than five college-level classes in mathematics (mean = 1.49). Students obtained partial course credit for participating.

Design & Procedure

Phase 1: defining a referent network

The first phase of the experiment required the three instructors to perform a concept relatedness task. All the instructors rated how closely related eight physics concepts were, using all pair-wise combinations of the concepts. Appendix A shows these 8 concepts. Both the order of presenting the concept pairs, and the left-right positioning in a pair, were made in a random order. A 5-point Likert scale (1= “Not at all related” and 5= “very related”) was adopted, while responses were pencil and paper based. The Pathfinder scaling technique was applied to these ratings data in order to reveal the underlying cognitive structures of the instructors. A referent network was constructed by averaging the relatedness ratings of the three instructors. Justification for this approach, discussed earlier, has been provided by the research of Acton, et al. (1994).

Phase 2: student performance

Student participation began in this phase of the experiment. The students were first asked to rate the relatedness of the 8 physics concepts in Appendix A. The procedure was the same as that described for the instructors in phase 1. Next, students were given time to study six solved physics word problems. Each of the problems was solved using two of the four equations in Appendix A. Half of the problems were from the subdomain of Kinematics while the other half were from the Dynamics subdomain. A kinematics problem, for example, might require students to solve for distance using the equations $\bar{v} = \frac{s}{t}$ and $a = \frac{\Delta v}{t}$. Likewise, a dynamics problem might require them to solve for acceleration using the equations $F = m \times a$ and $W = F \times s$.

Other examples of the problems and solutions are included in Appendix B. After this study phase, a second relatedness task was conducted with the same 8 concepts. Participants were not given access to the solved examples, equation sheet or the initial ratings during this exercise. Their background information on mathematics and physics was also collected as shown by the questionnaire in Appendix C. Finally, the participants were given six problems to solve. Two of the problems were identical to those given in the earlier task (referred to as *old transfer problems*). Two other problems had a different cover story but required the same equations in the same order for a correct solution (referred to as *near transfer problems*). Although the remaining two problems (referred to as *far transfer problems*) required the same equations as the other four, they had a different cover story and required a different order and manipulation of the original equations in the training examples. Appendix D shows some examples of these types of transfer problems.

Phase 3: scoring problem solving performance

The example below demonstrates how problem solving performance was scored in the present study. This scoring method is considerably more detailed than the method used in the original study. For instance, there was a deliberate attempt to reflect the knowledge required at each specific step in the solutions.

Example 1

A cat jumps from the roof of a building, uniformly accelerating downward at a rate of 10 m/s^2*, with an average velocity of 5 m/s. If the cat is travelling 2 m/s when it passes one window and 8 m/s when it passes another window, how far apart are the windows?*

change in velocity = 8 m/s - 2 m/s = 6 m/s

acceleration = $10\ m/s^2$

average velocity = 5 m/s

The definition of acceleration is given by:

acceleration = change in velocity / time (1) step 1

Rearranging terms, we find that:

time = change in velocity / acceleration (1) step 2

Now, substituting the change in velocity and acceleration of the cat:

time = $(6\ m/s) / (10\ m/s^2)$ (1) step 3

= .6 s (1) step 4

The definition of average velocity is given by:

average velocity = distance / time (1) step 5

Rearranging terms, we find that:

distance = average velocity x time (1) step 6

Finally, substituting the average velocity of the cat and the time calculated above:

distance = (5 m/s) x (.6 s) (1) step 7
= ***3 m*** (1) step 8

Total Possible Marks: (8)

All the problems were scored out of eight with a single mark given for each relevant step carried out:

Choose equation (step 1 & 5)

Manipulate equation (step 2 & 6)

Substitute values into equation (step3 &7)

Calculate numerical solution (step 4 & 8).

A step was marked correct if the correct value was obtained (regardless of units), if one of the values given in the problem statement was copied incorrectly (e.g. copying “2133” rather than “213.3”) but the rest of the solution was correct, or if the solution was correct except for an obvious arithmetic error that was likely due to incorrect values being entered into the calculator (e.g. if the participant wrote “333/33=100”). Appendix E shows a detailed score sheet for individual participants.

Phase 4: assessing structural knowledge

In addition, the present study employed Pathfinder in order to construct student PFnets from the second set of similarity ratings (i.e., following the studying of solved problems).

The resulting PFnets were compared to the referent PFnet in order to calculate similarities (i.e., PFsim values) as a measure of structural knowledge. In actuality, PFsim values were calculated at three different levels of specificity: overall (i.e., analyzing links among all 8 concepts), domain (i.e., analyzing links among the five concepts within a particular domain), and equation (i.e., analyzing only the links among the three concepts within a given equation). Figure 5, for example, shows the kinematics domain concepts as the five concepts joined by the blue and brown links in the referent network; consequently, the domain PFsim for the student is calculated based on the links among those five concepts. The "work" equation PFsim for example, is calculated based on the number of links among the three concepts with green links in the referent network.

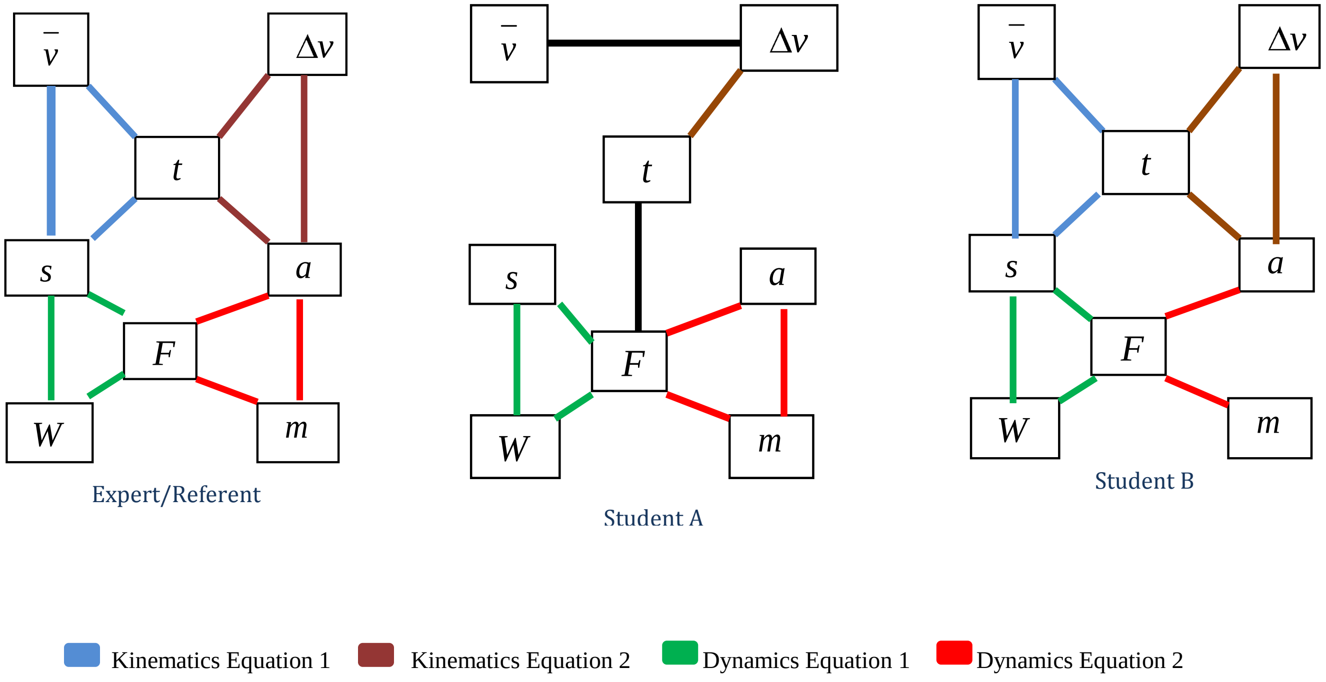

Figure 5. Example of hypothetical PFnets derived from equations in Appendix A.

Phase 5: correlational analysis

Overall Level

There are several ways to look at the diagnostic nature of Pathfinder networks. The first way to look at it is by how the overall similarity between students' networks and the expert network will predict overall performance on the physics problems. From the example in figure 5 above, it follows that Student B's overall performance is expected to be higher than that of Student A, based on PFSim. The latter only has 7 similar links to the referent and 2 irrelevant ones. In contrast, the former shares 11 of the 12 referent links, one missing link, and no irrelevant links.

This study conducted correlation analyses using the overall PF similarity coefficient as the independent variable while overall problem performance was the dependent variable. In this study, problem performance was assessed using three categories, i.e. old transfer, near transfer and far transfer problem performance, thereby yielding three unique outcome variables. Analyzing problem performance in this way helps to monitor and limit the influence of rote memorization and simple recall by participants since the three categories tap into different kinds of knowledge. It was expected that PF similarity would be a better predictor of far transfer problem solving than old or near transfer. In fact, structural knowledge, on which SAK is based, has been shown to be more related to far than near transfer (Trumpower, Guynn, & Goldsmith, 2003).

Domain Level

In addition to the overall similarity, a more diagnostic approach is a domain level similarity. We might expect similarities at the domain level to predict performance on specific types of problems and not others. In this case, similarity in PFnets on dynamics concepts (those shown in red & green in figure 5) as opposed to kinematics (blue & brown) concepts, would be expected to better predict performance on the dynamics problems only and not on the kinematics problems. Notwithstanding that figure 5 just showed Student B might be expected to do better overall, it also shows that Student A might be expected to do better than Student B on the dynamics problems. A closer look at the PFnets will reveal that Student A has exactly the same six links as the expert in that domain, while Student B has one missing link between *mass* and *acceleration* for the equation **F = m x a.**

For example, by looking at just the blue and brown portions of the PFnets, we could calculate the Kinematics PF similarities between the referent network and student networks. As at the overall level, the study conducted correlation analyses based on the three categories of domain problem performance, i.e. correlations between Kinematics PF similarity and old, near and far transfer Kinematics problem solving performance were derived. Likewise, correlations between Dynamics PF Similarity (red and green portion) and old, near and far transfer Dynamics problem solving performance were also derived.

Equation Level

An even more diagnostic approach proposed here involves analyzing similarities between links in a single equation within a domain. Do PF similarities based on only the concepts within an equation predict problem solving performance in problems that require knowledge of the conceptual relationships of that equation? There are two steps to solving the example above which require both Equations 1 and 2. The first step deals with “acceleration” and requires Kinematics Equation 2 to solve, while the second step deals with “distance” and requires Equation 1. We would expect a correct solution for step 1 from a participant who has the correct conceptual relationships of the Kinematics

Equation 2 (brown portion of Fig. 5) and Kinematics Equation 1 (blue portion of Fig. 5) for step 2, such as Student B. In order to test equation level similarities and their predictiveness of performance at the step level, the following table illustrates the types of analyses intended (see also attached Excel file).

EQUATION	ANALYSIS	STEP
Kinematics Equation 1 PF Similarity	correlation	old, near, far transfer kinematics step 1 performance
Kinematics Equation 2 PF Similarity	correlation	old, near, far transfer kinematics step 2 performance
Dynamics Equation 1 PF Similarity	correlation	old, near, far transfer dynamics step 1 performance
Dynamics Equation 2 PF Similarity	correlation	old, near, far transfer dynamics step 2 performance

Figure 6. Equation level analysis

Figure 6 shows that the problem solving performance on each of the three transfer categories will be correlated with the PFsim for that equation.

VI. RESULTS

Pathfinder networks for the students and the experts were generated and analyzed for similarities. Figure 7, below, shows the pathfinder network that was generated for the averaged expert relatedness ratings.

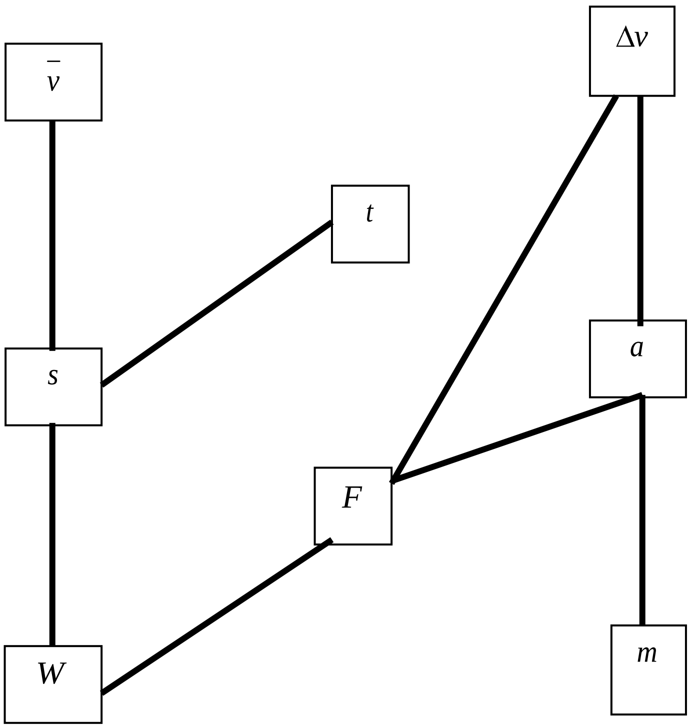

Figure 7. Pathfinder generated PFnet based on averaged expert ratings.

Recall that the expert network provides a standard against which the students' networks can be evaluated. A visual inspection of the networks revealed similarities and differences in the links that characterize the expert knowledge structure and the structures for the students. When the students' PFnets were evaluated against the average expert PFnet using the set-theoretic formula (intersection/union), corresponding PFsims for each student were attained at each level of specificity described earlier. Table 1 below shows the descriptive statistics for the calculated similarity indices (PFSims).

Table 1

Mean PFSim at Each Level of Analysis

Level of Analysis	N	Mean	SD
Overall	47	.24	.08
Kinematics	47	.38	.21
Dynamics	47	.42	.20
Kinematics Equation 1	47	.44	.34
Kinematics Equation 2	47	.33	.29
Dynamics Equation 1	47	.50	.26
Dynamics Equation 2	47	.35	.24

The mean student - expert PF similarities are consistent with the literature. For example, Goldsmith et al (1991) suggested that a student – expert PFsim of between .3 and .5 is fairly good. Expert – expert similarities are usually around .8. However, in some cases the std. deviations were relatively high indicating that the scores were widely distributed.

Table 2

Mean Problem Solving Performance at Each Level of Analysis

Level of Analysis	N	Mean	SD
Overall	47	29.66	8.96
Old Problems	47	6.17	6.40
Near Transfer Problems	47	12.74	6.97
Far Transfer Problems	47	10.74	4.36

Kinematics	47	12.66	6.02
Old Problems	47	2.94	3.53
Near Transfer Problems	47	5.51	4.23
Far Transfer Problems	47	4.21	3.09
Dynamics	47	17.00	5.38
Old Problems	47	3.23	3.53
Near Transfer Problems	47	7.23	4.56
Far Transfer Problems	47	6.53	2.49
Kinematics Equation 1	47	6.32	3.23
Old Problems	47	1.49	1.85
Near Transfer Problems	47	2.70	2.24
Far Transfer Problems	47	2.13	1.74
Kinematics Equation 2	47	6.34	3.47
Old Problems	47	1.45	1.83
Near Transfer Problems	47	2.81	2.32
Far Transfer Problems	47	2.09	1.79
Dynamics Equation 1	47	8.83	3.05
Old Problems	47	2.00	2.25
Near Transfer Problems	47	3.72	2.52
Far Transfer Problems	47	3.26	1.33
Dynamics Equation 2	47	8.40	2.86
Old Problems	47	1.62	1.84

Near Transfer Problems	47	3.51	2.34
Far Transfer Problems	47	3.28	1.39

Solutions to each problem on the problem solving task were scored out of eight in accordance with the scoring method proposed in the methodology. In some cases, participants skipped some of the steps to solving the problem but still managed to reach the correct solution. In these cases, full marks were awarded because while they did not explicitly state the algebraic manipulations on paper, it was unlikely that they could generate a correct numerical answer without an implicit understanding of the solution. Participants' problem solving performance was also scored at each level of specificity (i.e., across all problems, for only the problems in a particular domain, and for the steps within problems that required the use of a given equation). The maximum possible scores for each level of analysis were 48, 24 and 12 for the overall, domain and equation levels, respectively. Further, each of these analysis level scores can be divided by three to get the maximum possible score for the old, near and far transfer components. Table 2 shows descriptive statistics of the resulting scores on the problem solving task. Appendix E shows the complete table for individual scores.

In order to assess the predictive capabilities of SAK, the study conducted several Pearson correlations between the PF similarities of the students and their respective problem solving scores. The columns in table 3 below show the results of the relationships between the data in tables 1 and 2.

Table 3

Pearson Correlations between PFsim and Problem Solving Performance at Each Level of Analysis

Level of Analysis	r
Overall PFsim-Overall Problems	.525***
Overall PFsim-Overall Far Transfer Problems	.550***
Overall PFsim-Overall Near Transfer Problems	.275*
Overall PFsim-Overall Old Problems	.061
Kinematics PFsim-Kinematics Problems	.248*
Kinematics PFsim-Kinematics Far Transfer Problems	.217^
Kinematics PFsim-Kinematics Near Transfer Problems	.135
Kinematics PFsim-Kinematics Old Problems	.072
Dynamics PFsim-Dynamics Problems	.205^
Dynamics PFsim-Dynamics Far Transfer Problems	.197^
Dynamics PFsim-Dynamics Near Transfer Problems	.043
Dynamics PFsim-Dynamics Old Problems	.117
Kinematics Equation 1 PFsim-Kinematics Problems Equation 1	.081
Kinematics Equation 2 PFsim-Kinematics Problems Equation 2	.118
Dynamics Equation 1 PFsim-Dynamics Problems Equation 1	.126
Dynamics Equation 2 PFsim-Dynamics Problems Equation 2	.155

^p<.10, one-tailed, *p< .05, one-tailed, **p< .01, one-tailed, ***p< .001, one-tailed

After considering the inflated likelihood of making a type I error (often referred to as a ’false positive’, and is the process of incorrectly rejecting the null hypothesis in favor of the alternative) as a result of conducting numerous significance tests, the main analysis was limited to the seven correlations between PFsim and problem solving performance at the same level of analysis that directly relate to the goals of the study (e.g., the correlation between PFsim based solely on the kinematics concepts and problem solving performance on just the kinematics problems). Each of the above that was significant was then followed-up with the correlations between the PFsim and the old, near, and far transfer problem solving performance at that particular level. The alpha level used for the seven main correlations was .05, but the alpha level used for follow-up tests was a Bonferonni corrected .05/3=.017.

As previous studies have indicated, the current results confirm that overall PFsim is a good measure of overall knowledge. The correlation coefficient (r=.525) shows a relatively strong relationship between PFsim and problem solving performance at the global levels. Additionally, and as hypothesized, follow-up correlations reveal that the relatively strongest relationship between the two variables is observed when overall PFsim is compared to overall *far transfer* problem solving (r=.550). Neither the correlation with near transfer nor old problems was significant.

At the domain level, PFsim was significantly correlated with problem solving performance for the kinematics domain (r=.395), but only marginally for the dynamics domain (r=.205). The pattern of the follow-up correlations, despite a lack of significance after the Bonferroni correction, indicate a general tendency for PFsim to positively relate more strongly to far transfer performance than to near transfer or old problem performance. This finding is similar to results at the overall level.

At the equation level, none of the correlations were statistically significant. A possible explanation for the lack of significance could simply be due to error since the calculation for PFsim at this extremely specific level is not stable. A lot of variance is introduced into these calculations as the level of analysis becomes more and more specific because PFsim is obtained from fewer and fewer links. As discussed later, this could be a signal to the extent with which the specificity of SAK can be utilized.

VII. DISCUSSION

The current study sheds further light on SAK (and Pathfinder Networks) – its capabilities and what it measures. The following three sections will each discuss some of the major issues that were the inquiry of the study. Section (a) deals with issues of student conceptions and how the results have managed to address this issue. Section (b) addresses issues of competence and validity, among other issues. It also draws attention to how the results in this study address these particular issues. Finally, Section (c) is an extension of Section (b), although it deals with competence at the extreme levels of analysis, such as at the equation level. This section highlights the difficulties associated with extra specification of performance levels.

a. Alternative conceptions

The results attained here support the earlier prognosis that the alternative conceptions held by various students will translate into the specific types of errors they make when attempting procedural transfer tasks such as problem solving. If we consider, as past research has shown, the expert network as a close approximation to a domain's knowledge structure, then any missing or irrelevant links in the student's network when evaluated against the expert network would represent alternative conceptions. It follows, therefore, that the errors in solving problems whether at the domain level, or indeed at the equation level, can be attributed to such alternative conceptions. Earlier we saw that SAK has been employed in previous studies for purposes of establishing overall structural knowledge of a learning domain, and this was shown to be valid by, among many others, Trumpower and Goldsmith (2004); but the individual nodes generated by pathfinder networks have rarely been used to study alternative conceptions or weak conceptions surrounding sub-domains or individual concepts. As expected, the present study also confirmed the earlier findings by way of a statistically significant correlation between the overall measures of PF similarity and problem solving. The general implication of this finding is that, students with poor structural knowledge (or alternative conceptions) of a discipline as assessed by SAK perform poorly on procedural tasks within that discipline and vice-versa (Day, Arthur, & Gettman, 2001; Kraiger, 1993; Trumpower & Goldsmith, 2004), thereby indicating the predictive ability of SAK and the importance of structural knowledge. Nevertheless, the findings on overall network similarity measures only indicate *how much* students' structures differ from experts' structures at a global level; they do not indicate specifically *in which sub-domain* structures differ. The quest for an answer to this question was one of the major motivations for the current study.

b. Domain Competence

The specificity of SAK was validated and given impetus by the correlation results attained in section VI above. The results showed that PFsim was more related to performance on far transfer problems than any other type of problems. Out of three, there was only one statistically significant relationship between near transfer problems and

PFsim at both the overall and sub-domain levels of analysis. None of the old transfer problems attained a statistically significant relationship with PFsim at these same levels. Interpretation of these findings lead to the conclusion that far transfer problem performance is an indicator of a deeper level understanding of conceptual relationships, and therefore conceptual or structural knowledge. Recall that old transfer problems were the exact same problems initially given to participants as study examples and then tested in the problem solving task. It would not be far-fetched to assume the effect of rote memorization and simple recall when participants correctly solve these problems. Also recall that, Ausubel (1968) managed to distinguish between meaningful and rote learning, and rejected the verbatim repetition of definitions and the reproduction of procedures as meaningful learning. Due to this consideration, it was hypothesized that performance on these problems would not necessarily reflect deeper measures of structural knowledge. Unsurprisingly, old transfer problem performance did not exhibit any statistically significant relationship with corresponding PFsims. These conclusions provide evidence for the existence of convergent and discriminant validity for the measures used in the study. We have convergent validity when measures of constructs that theoretically *should* be related to each other are, in fact, observed to be related to each other (Rupp et al, 2010), as in the case of far transfer performance and pathfinder networks in the present study. Discriminant validity is observed between the mostly low and statistically insignificant relationships between old or near transfer performance and Pathfinder similarity indices. Discriminant validity is present when measures of constructs that theoretically should *not* be related to each other are, in fact, observed to not be related to each other (Rupp et al, 2010). Since it is clear that Pathfinder networks tap into an individual's conceptual (or structural) knowledge, and old transfer performance does not, then a correlational relationship should not exist between these measures. To further illustrate this point, correlations across domains (e.g., the correlation between the PFsim for dynamics and the problem solving performance for kinematics and vice versa) were also conducted. The correlation between PFsim based on kinematics concepts and problem solving performance on far transfer dynamics problems was not statistically significant, r=.002, p>.100 (one-tailed), nor was the correlation between PFsim based on dynamics concepts and problem solving performance on far transfer kinematics problems, r=.023, p>.100 (one-tailed). Comparing these results to those in Table 3, it is clear that the current Pearson values are much lower and statistically insignificant. Thus, there is even more discriminant validity evidence within the study – that is, the results show that it is not just overall conceptual knowledge that is responsible for performance, but rather, it is the conceptual knowledge within the domain that is responsible for the observed far transfer problem solving performance. This is an interesting and important finding which further buttresses the specificity hypothesis of SAK.

Earlier, it was postulated that with an alternative and more fine-grained approach to the PFnets derived from SAK, students' specific strengths and weaknesses could be identified. In particular, the presence of *specific* links in students' PFnets could be associated with their ability in solving *specific* types of problems. The present findings suggest that subsets of links in PFnets can predict the presence of specific strengths or weaknesses. The problem solving performance in a sub-domain by participants who possessed more of that link-subset was relatively stronger than of those who did not have as many of these links. These findings indicate that links in PFnets represent specific bits of structural knowledge that have particular consequences when attempting to apply one's knowledge. Thus, it would appear that a fine-grained evaluation of links within students' PFnets can be used to identify specific areas of weakness to be targeted in further instruction.

c. Further Specificity

To this point, the discussion of the results has been mainly limited to the global and domain levels only. The third stage of the analysis was at the equation level. Earlier in the study, the question was asked: Do PF similarities within an equation predict problem solving performance in problems that require knowledge of the conceptual relationships of that equation? At that time, it was postulated that, a correct solution to either step 1 or 2 of a problem would require knowledge of the conceptual relationships constituting the appropriate subdomain equation. Table 3 above shows these equation level relationships between PFsim and problem solving. Contrary to expectation, there was no significance in the relationships at these levels. One possible explanation for the lack of finding significance could just be that it is not possible to define conceptual understanding with this level of specificity. It may be more appropriate to conclude that the one significant relationship attained is attributable to pure chance than to any meaningful explanation for conceptual knowledge at this extremely specific level of analysis. This finding could signal the existence of demarcation points or limits to the specificity with which we can employ SAK. Proof of this deduction and conclusion will be left for further investigation and future research.

Another valid explanation alluded to earlier, is more of a mathematical rather than a conceptual explanation for why the results were not significant at the most specific level. PFsim is less stable (and, therefore, less related to problem solving performance) at this level due to the calculation of its value. Whereas overall PFsim is calculated based on the twelve links in Figure 5, the domain PFsim and the equation PFsim are based on six and three links, respectively. In other words, one link represents about 8%, 17% and 33%, at the overall, domain and equation levels, respectively. Therefore, the more specific the analysis gets, the more variance (or instability) is introduced by a unit change in the number of links present in a PFnet. Therefore, if PFsim is not stable (or reliable), then its relationship with problem solving performance cannot be expected to be reliable either. This implies that it is not possible to make reasonable predictions about performance based on this measure. As a result, the results at the very specific analysis levels need to be taken with caution.

VIII. LIMITATIONS

Although this study focused on Physics in analyzing alternative conceptions and their impact on problem solving performance, the same logic that guided this study holds true for all scientific disciplines such as Mathematics, Biology or Chemistry. The available secondary data dictated the discipline (i.e. physics) being studied here. Nevertheless, this constraint was not material in the formulation of the methodology for this research because SAK easily adapts to many disciplines. It has previously been used in many other disciplines including author co-citation analysis (White, 2003), flight training

(Schvaneveldt, Beringer, & Lamonica, 2001), and software requirement understanding (Kudikyala & Vaughn, 2005).

The limitations to this research were ultimately defined by its quality and its context. Notwithstanding that rigorous data analysis strengthened the quality of the study and limited concerns about validity issues each individual study is also limited by its own contextual nature. The implications of this study, therefore, are constrained by the limitations of SAK. One of the weaknesses of the SAK method is that it is quite tedious and is useful for a limited size set of concepts. The number of pair-wise comparisons that a participant rates is equal to n (n-1)/2, where n is the total number of concepts in the set. The number of pair-wise comparisons increases exponentially with each concept added to the set. So, for example, a set of 13 concepts results in 78 pair-wise comparisons, but a set of 26 concepts results in 325 pair-wise comparisons. Such a big set would require a substantial amount of time for the students to rate, and this time is not always available in the classroom. Students can also become fatigued, distracted or bored with too many ratings. The instructors or researchers therefore choose a limited number of concepts to keep the rating task manageable. Goldsmith, Johnson, and Acton (1991) argued that “the decrease in predictive validity with smaller sizes is a fairly linear function”, therefore it seems reasonable to include only important concepts in the set.

A shortcoming of Pathfinder networks is also that, while they are able to identify areas of weakness (i.e., alternative conceptions), they cannot yet identify the source of those alternative conceptions. For example, a missing link might indicate a problem, but it does not diagnose *why* the student does not have the link. Pathfinder is not able to tell us what the exact cause of the missing link is – but just that it exists.

Pathfinder is also limited by the exact concepts chosen for a particular ratings task. For example, a physics expert's conceptual structure likely has links to many other concepts in addition to the 8 used in this study. Put differently, the validity of PFsim as a measure of conceptual knowledge is limited by the validity of the set of concepts chosen to assess the knowledge.

IX. SUMMARY

Several methods exist for assessing mental models which highlight an individual's conceptions about a domain. Some of these measures are characterized by weaknesses, such as the subjectivity inherent in measurement approaches like concept mapping. SAK was used here, and as shown before, is one of the ways to measure knowledge of conceptual relationships and associated alternative conceptions. Recall it was hypothesized that:

1. Students' PFsims would predict their problem solving performance.
2. Far Transfer problems require higher levels of conceptual understanding; consequently PFsims would have stronger positive correlation with far transfer than old or near transfer problem solving performance.
3. The presence or absence of specific links in students' PFnets would predict specific alternative conceptions held by the students. These alternative conceptions would manifest in the specific types of problems solved correctly or incorrectly.

As the results have shown, the present study managed to test and confirm these hypotheses, although marginally under some circumstances. By employing pathfinder networks, SAK seems potentially useful as an alternative to other assessment methods, because it easily transfers from one learning domain to another. This study explored how predictive of performance this method was at the more specific levels of network analysis. PFnet analysis has more commonly been employed to acquire more general conceptual relations. However, the present study expanded the approach by looking at closer, more specific relationships at the individual network links level. The attained results seemed to support what had initially been hypothesized, despite the positive associations being generally weaker at the domain and especially equation levels. These findings can therefore represent a contribution to student assessment, learning and instruction in general.

More specifically, the results at the overall level gave support to the usefulness of SAK as a tool for summative assessments. The overall PFsim was highly positively correlated to overall student performance in problem solving. Subsequently, all performance was determined by this single PFsim score as in typical summative assessments. Therefore SAK can be used summatively at the end of a course or a unit to assign a type of grade, based on the overall similarity. Using it after every lesson might be difficult because of few concepts and the accompanying instability problems discussed earlier. SAK could also be used at the beginning of a unit to guide instruction (assuming students come in with prior knowledge). Students' PFnets at the beginning of a learning period could help to identify teaching areas that need emphasis. Instructors have the benefit to tailor instruction according to the identified weaknesses or strengths of the class. During instruction, SAK could also be used to show students the areas they need to concentrate on for unit tests, reviews, etc. Presenting students with their PFnets before some type of test could help them prepare adequately.

One of the things this study also explored was whether or not SAK satisfied a primary condition of a formative assessment tool – the ability to identify specific areas of strength and weakness. Since hypothesis 3 was satisfied, at least at the overall and domain levels (the presence or absence of specific subsets of links was able to predict performance on problem solving), it seems reasonable to conclude that the first condition for formative assessment was also established for SAK. It follows, therefore, that further investigation into SAK's ability to satisfy subsequent conditions of effective formative assessment tools (e.g., that they generate useful feedback) should be conducted.

In conclusion, SAK should be adopted as an alternative summative assessment tool. This is not a new finding, but it supports previous studies conducted at the global level. Present findings suggest that SAK is reasonably capable of predicting performance on levels that are more specific than just the overall level. In that sense, SAK can be adopted as an alternative "diagnostic" assessment tool for specific sub-domains in many disciplines. The results obtained at the equation level however suggest that further research on the extent of specificity should also be conducted. The results show instability as the analysis gets more specific.

X. REFERENCES

Acton, W.H., Johnson, P.J., & Goldsmith, T.E. (1994). Structural knowledge assessment: Comparison of referent structures. *Journal of Educational Psychology, 86*(2), 303–311.

Adelson , B. (1981). Problem solving and the development of abstract categories in programming languages. *Memory and Cognition.* 9(4), 422-433.

American Association for the Advancement of Science. (2000). http://www.project2061.org/research/textbook/hsbio/derault.him

Anderson, J.R. (I983). Problem solving and learning. *American Psychologist, 48.* 35-44.

Anderson, J. R. (1995). *Cognitive psychology and its implications*. New York: Freeman.

Anderson, J. R., Bothell, D. Byrne, M.D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036-1060.

Anderson, D., Lucas, K. B., Ginns, I. S., & Dierking , L. D. (2000). Development of knowledge about electricity and magnetism during a visit to a science museum and post-visit activities. *Science Education, 84,* 658-679.

Ausubel , D.R (1963). *The psychology of meaningful verbal learning.* New York: Grune & Stratton.

Ausubel, D.R (1968). *Educational psychology: A cognitive view.* New York: Rinehart and Winston.

Baxter, G.P., Elder, A.D., & Glaser, R. (1996). Knowledge-based cognition and performance assessment in the science classroom. *Educational Psychologist, 31,* 133-140.

Black, P. (2004). Raising standards through formative assessment. A paper presented to the General Teaching Council for England (GTC) conference, London. (p.16).

Boo, H., & Watson, J. R. (2001). Progression in high school students' (aged 16-18) conceptualizations about chemical reactions in solution. *Science Education, 85,* 568-585.

Borgen, K. L., & Manu, S. S. (2002). What do students really understand? *Journal of Mathematical Behavior,21,* 151-165.

Bransford, J. D., Brown, A. L., & Codking, R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, D.C.: National Academies Press.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: Bradford.

Chi, M.T.H. (1993). Barriers to conceptual change in learning science concepts: A theoretical conjecture, ln W. Kintsch (Ed.), *Proceedings of the Fifteenth Annual Cognitive Science Society Conference* (pp. 312-317). Hillsdale, NJ: Erlbaum.

Chi, M.T.H., Slotta, J.D., & deLeeuw, N. (1994). From things to proces.ses: A theory of conceptual change for learning science concepts. *Learning and Instruction. 4.* 27-43.

Costu, B., & Ayas, A. (2005). Evaporation in different liquids: Secondary students' conceptions. *Research in Science & Technological Education, 23*(1), 75-97.

Creswell, J. W. (2003). *Research Design:Qualitative, Quantitative, and Mixed Methods Approaches* (2^{nd} Eds.). Thousand Oaks, CA: Sage Publications.

d'Apollonia, S., De Simone, C. Dedic, H., Rosenfieid, S., & Glashan, A. (1993). *Cooperative networking: A method of promoting understanding in the sciences.* Final Report submitted to Programme d'Aide a la rescherche sur I'enseignement et I'apprentissage (PAREA). Gouvenement du Quebec. ISBN 0-969728-1-0.

Dayton, T., Durso, F.T., & Shepard, J.D. (1990). A measure of the knowledge reorganization underlying insight. In R.W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization* (pp. 267–277). Norwood, NJ: Ablex.

DiCerbo, K.E. (2007), Knowledge structures of entering computer networking students and their instructors, *Journal of Information Technology Education, 6*, 263-277.

Driver, R., & Easley, J. (1978). Pupils and paradigms: A review of literature related to concept development in adolescent science students. *Studies in Science Education, 5,* 61-84.

Earl, L. (2003). *Assessment as learning: Using Classroom Assessment to Maximise Student Learning.* Thousand Oaks: Corwin Press.

Egan, D.E., & Schwartz. B.J. (1979). Chunking in recall of symbolic drawings. *Memory and Cognition,* 7. 149-158.

Fenker, R.M. (1975). The organization of conceptual materials: A methodology for measuring ideal and actual cognitive structures. *Instructional Science,* 4 (1), 33-57.

Ferrari. M.. & Chi, M.T.H. (1998). The nature of naive explanations of natural selection. *International Journal of Science Education, 20.*, 1231-1256.

Frederickson, C.H., & Breaueux, A. (1990), Applying cognitive task analysis and research methods to assessment. In N. Frederickson, R. Glaser., A. Lesgold. & M.G. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (pp. 143-165). Hillsdale. NJ: Lawrence Erlbaum.

Gardner, P.L. (1986). Physics students' comprehension of motion with constant velocity. *The Australian Science Teachers Journal, 31*. 27-32.

Gentner, D., & Stevens, A. (Eds.). (1983). *Mental models.* Hillsdale. NJ: Erlbaum.

Gilbert, J.K., & Watts, D.M. (1983). Concepts, misconceptions and alternative conceptions: changing perspectives in science education. *Studies in Science Education,10*, 61-98.

Goldsmith, T.E., & Davenport, D.M. (1990). Assessing structural similarity of graphs. In R.W. Schvaneveldt (Ed.), Pathfinder associative networks: Studies in knowledge organization (pp. 7587). Norwood, NJ: Ablex.

Goldsmith, T.E., Johnson, P.J., & Acton, W.H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 35* (1),88-96.

Goldsmith, T.E., & Kraiger, K. (1997). Applications of Structural Knowledge Assessment to Training Evaluation. In J. Kevin Ford, *Improving Training Effectiveness in Work Organizations* (pp. 73 – 96). Mahwah, NJ: Lawrence Erlbaum Associates.

Gouli, E., Gogoulou, A. & Grigoriadou, M. (2003). A coherent and integrated framework using concept maps for various educational assessment functions. *Journal of Information Technology Education, 2*, 215-240. Retrieved May 10, 2007 from http://www.jite.org/documents/Vol2/v2p215-240-23.pdf

Griffiths. A.K.. & Grant, B.A.C. (1985). High school students' understanding of food webs: Identification of a learning hierarchy and related misconceptions. *Journal of Research in Science Teaching. 22,* 421 -436.

Guisasola, J., Almudi, J. M., & Zubimendi, J. L. (2004). Difficulties in learning the introductory magnetic field theory in the first years of university. *Science Education, 88*(3), 443-464.

Guskey, T.R. (2003). How Classroom Assessments Improve Learning. *Educational Leadership*, 60(5), 1-6.

Hackling, M.W., & Garnet, P.G. (1986). Chemical equilibrium: Learning difficulties and teaching strategies. *The Australian Science Teachers Journal,* i7, 8-13.

Hailikari, T., Negri, A., & Lindblom-Ylanne, S. (2007). Exploring alternative ways of assessing prior knowledge, its components and their relation to student achievement: A mathematics based case study. *Studies in Educational Evaluation*, 33 (3/4), 320-337.

Halhoun, I.A., & Hestenes. D. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, 5, 1056-1065.

Hamrick, L., Hearty, H., & Ault, C. (1987). Concept structure interrelatedness competence (ConSic): A tool for examining and promoting cognitive structure. *School Science & Mathematics*, 87,655-664.

Hammer, D. (1996). More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research. *American Journal of Physics*, *64*(10), 1316-1325.

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

Jacobson. M.J., & Archodidou, A. (2000). The design of hypermedia tools for learning: Fostering conceptual change and transfer of complex scientific knowledge. *The Journal of the Learning Sciences*, 9(2). 149-199.

Jee, B.D., & Wiley, J. (2007). How goals affect the organization and use of domain knowledge. *Memory & Cognition*, 35 (5), 837-851.

Johnson, P.E., (1964). Associative meeting of concepts in physics. *Journal of Educational Psychology*, 55, 84-88.

Johnson, P.E., Cox, D.L., & Curran, T.E. (1970). Psychological reality of physical concepts. Psychonomic Science, 19, 245-247.

Johnson, P.J., Goldsmith, T.E., & Teague, K. (1995) Similarity, structure, and knowledge: A representational approach to assessment. In P.D. Nicholas, S.F. Chipman, R.L. Brennan, (Eds.), *Cognitively diagnostic assessment* (pp. 221-249). Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P.N. (1983). *Mental models.* Cambridge. MA: Harvard University Press.

Jonassen, J. H. & Grabowski, B. L. (1993). *Handbook of individual differences, learning and instruction. Part VII, Prior knowledge.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Khazanov, L. (2006). An investigation of approaches and strategies for resolving students'misconceptions about probability in introductory college statistics (Doctoral dissertation, Columbia University Teachers' College). *Dissertation Abstracts International, 66*, 10A, 3593.

Kinchin, I., Hay, D. B., & Adams, A. (2000). How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. *Educational Research, 42*, 43-57.

Korner, C. (2005). Concepts and misconceptions in comprehension of hierarchical graphs. *Learning and Instruction, 15*(4), 281-296.

Koubek, R.J., & Mountjoy, D.N. (1991). *Toward a model of knowledge structure and a comparative analysis of knowledge structure measurement techniques.* ERIC_NO:ED339719

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, *78*(2), 311-328.

Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. *Psychometrika*, 29, 115129.

Kudikyala, U.K. & Vaughn, R.B. (2005). Software requirement understanding using pathfinder networks: Discovering and evaluating mental models. *The Journal of Systems and Software*, 74 (1), 101-108.

Mahon, B.Z., & Caramazza, A. (2007). The organization and representation of conceptual knowledge in the brain: Living kinds and artifacts. In E. Margolis & S. Laurence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 157-187). NY: Oxford University Press.

Matsumura, M. (1998). Is it fair to teach evolution? *Reports of the National Center for Science Education. 18.* 19-21.

McClure, J., Sonak, B., & Suen, H. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching, 36*, 475-492.

McComas, W.F. (1997). The discovery and nature of evolution by natural selection: Misconceptions and lessons from the history of science. *American Biology Teacher.59*, 492-500.

McManus, S. (2006). Attributes of Effective Formative Assessment. Retrieved July 31, 2009 from The Council of Chief State School Officers website: http://www.ncpublicschools.org/docs/accountabilty/educators/fastattributes04081.pdf.

McNamara. TP., Miller. D.L.,& Bransford, J.D. (1991). Mental models and reading comprehension. In R. Bait, M. Kamil, P. Mosenthal, & PD. Pearson (Eds.). *Handbook of research in reading* (pp. 490-511). New York, NY: Longman.

Mogharreban, N. (2004). Approximate degrees of similarity between a user's knowledge and the tutorial systems' knowledge base. *Journal of Information Technology Education, 3,* 219-226. Retrieved May 10, 2007 from http://www.jite.org/documents/Vol3/v3p219-226-138.pdf

Mosenthal, P.B.. & Kirsch, I.S. (1992). Learning from exposition: Understanding knowledge acquisition from a knowledge model perspective. *Journal of Reading, 35,* 356-370.

Murphy. G.L.. & Wright. J.C. (1984), Changes in conceptual structure with expertise: Differences between real-world experts and novices. *Journal of Experimental Psychology: Learning. Memory. and Cognition, 10,* 144—155.

Novak, J.D. (1988), Learning science and the science of learning. *Studies in Science Education,* 15. 77-101.

Novak, J. D. (1990a). Concept maps and vee diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional Science*, *19*(1), 29-52.

Novak, J. D. (1990b). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*(10), 937-950.

Novak, J.D., & Musonada, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal 28,* 117-153.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*, 937-950.

Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier Diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching, 32*, 45–61.

Olson, J.R., & Bioisi, K.J. (1991). Technique for representing expert knowledge. In K.A. Ericsson &J. Smith (Eds.), *Toward a general theory of expertise* (pp. 240-285). Cambridge: Cambridge University Press.

Piaget, J. (1954). *The construction of reality in the child.* New York: Basic Books.

Pine, K., Messer, D., & St. John, K. (2001). Children's misconceptions in primary science: A survey of teachers' views. *Research in Science & Technological Education, 19*(1), 79-96.

Preece, P.F.W, (1976). The concepts of electromagnetism: A study of the internal representation of the external structures. *Journal of Research in Science Teaching,* 13, 517-524.

Resnick, L. (1987). *Education and learning to think.* Washington, D.C.: National Academy Press.

Rice, D., Ryan, J., & Samson, S. (1998). Using concept maps to assess student learning in the science classroom: Must different methods compete? *Journal of Research in Science Teaching, 35*, 1103-1127.

Ruiz-Primo, M. A. & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33*, 569-600.

Rumelhart, D.E. (1980). Schemata: The building blocks of cognition. In RJ. Spiro. B.C. Bruce. & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.

Rupp, A. A., Henson, R.A., Templin, J. (2010). *Diagnostic measurement: theory, methods & applications.* New York: Guilford Press.

Schvaneveldt, R.W., Beringer, D.B., & Lamonica, J.A. (2001). Priority and organizations of information accessed by pilots in various phases of flight. *International Journal of Aviation Psychology, 11 (3)*, 253-280.

Schvaneveldt, R.W., Durso, F.T., & Dearholt, D.W. (1989) Network structures in proximity data. In G.H. Bower (Ed.) *The psychology of learning and motivation* (Vol. 24, pp. 249-284). New York: Academic Press.

Schvaneveldt, R.W. (1990). *Pathfinder associative networks: Studies in knowledge organization.* Norwood, NJ: Ablex.

Sencar, S., & Eryilmaz, A. (2004). Factors mediating the effect of gender on ninth-grade Turkish students' misconceptions concerning electric circuits. *Journal of Research in Science Teaching, 41*(6), 603-616.

Shaffer, D. M., & McBeath, M. K. (2005). Naive beliefs in baseball: Systematic distortion in perceived time of apex for fly balls. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*(6), 1492-1501.

Shavelson, R.J.,Ruiz-Primo, M.A., & Wiley, E. (2005). Windows into the mind. *International Journal of Higher Education,49*, 413 – 430.

Shavelson, R.J., (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology*, 63, 225-234.

Shavelson, R.J., & Stanton, G.C. (1975). Concept validation: Methodology and application to three measures of cognitive structure. *Journal of Educational Measurement*, 12, 67-85.

Shoenfeld, A.H.. & Herrmann. D.J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, H*, 484—494.

Slotta, J.D., & Chi, M.T.H. (1999). *Overcoming robust misconceptions through ontological training*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal.

Tallman, S., & Phene, A. (2007). Leveraging knowledge across geographic boundaries. Organization Science, 18 (2), 252-260.

Taylor, A. K., & Kowalski, P. (2004). Naïve psychological science: The prevalence, strength, and sources of misconceptions. Psychological Record, 54, 15-25.

Thornton, R. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws, The Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics*, *66*, 338-352.

Trumpower, D. L. (2003). *Development of problem solving performance and structural knowledge in physics problem solvers.* Doctoral dissertation. Department of psychology, University of New Mexico, Albuquerque.

Trumpower, D. L., & Goldsmith, T.E. (2004). Structural enhancement of learning. *Contemporary Educational Psychology, 29*(4), 426-446.

Trumpower, D. L., Guynn, M.J.,& Goldsmith, T.E. (2004). Goal specificity and knowledge acquisition in statistics problem solving: Evidence for attentional focus. *Memory & Cognition, 32* (8*)*, 1379-1388.

Trumpower, D.L., Sharara, H., & Goldsmith, Timothy, E (2010). Specificity of Structural Assessment of Knowledge. *Journal of Technology, Learning, and Assessment, 8* (5), 1-33.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4,* 45-69.

Vosniadou, S.. & Brewer, W.F. (1987). Theories of knowledge restructuring in development. *Review of Educational Research. 57,* 51-67.

Walker, J. M. T., & King, P. H. (2002). Concept mapping as a form of student assessment and instruction. *Proceedings of the 2002 American Society for Engineering Education Annual Conference.* Retrieved January 24, 2007 from http://www.vanth.org/docs/BMES_005.pdf

White, H.D. (2003). Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists. *Journal of American society for information science and technology, 54* (5), 423-434.

Williams, M. (2004). Concept mapping: A strategy for assessment. *Nursing Standard, 19*(9), 33-38.

Zaim-Idrissi, K.,Desautels, J., & Larochelle, M. (1993). The map is the territory! The viewpoints of biology students on the theory of evolution. *The Alberta Journal of Educational Research, 39,* 59-72.

APPENDIX A

Equations Used to Solve Problems in Experiments 2 - 4

Kinematics

$$\bar{v} = \frac{s}{t}$$

$$a = \frac{\Delta v}{t}$$

Dynamics

$$F = m \times a$$

$$W = F \times s$$

where, a =acceleration, $\bar{v}$ = average velocity, Δv = change in velocity, s = distance, F = force, m = mass, t = time, and, W = work

APPENDIX B

Solved Example Problems From Each Subdomain

Kinematics

A cat jumps from the roof of a building, uniformly accelerating downward at a rate of 10 m/s^2, with an average velocity of 5 m/s. If the cat is travelling 2 m/s when it passes one window and 8 m/s when it passes another window, how far apart are the windows?

change in velocity = 8 m/s - 2 m/s = 6 m/s

acceleration = 10 m/s^2

average velocity = 5 m/s

The definition of acceleration is given by:

acceleration = change in velocity / time

Rearranging terms, we find that:

time = change in velocity / acceleration

Now, substituting the change in velocity and acceleration of the cat:

time = (6 m/s) / (10 m/s^2)

= .6 s

The definition of average velocity is given by:

average velocity = distance / time

Rearranging terms, we find that:

distance = average velocity x time

Finally, substituting the average velocity of the cat and the time calculated above:

distance = (5 m/s) x (.6 s)

= **3 m**

Dynamics

A weightlifter performs 1950 Nm of work when he lifts a 130 kg barbell a distance of 1.5 m. What is the acceleration of the barbell?

mass = 130 kg

distance = 1.5 m

work = 1950 Nm

The definition of work is given by:

work = force x distance

Rearranging terms, we find that:

force = work / distance

Now, substituting the work performed and distance lifted:

force = (1950 Nm) / (1.5 m)

= 1300 N

The definition of force is given by:

force = mass x acceleration

Rearranging terms, we find that:

acceleration = force / mass

Finally, substituting the mass of the barbell and the force calculated above:

acceleration = (1300 N) / (130 kg)

= **10 m/s^2**

APPENDIX C

Questionnaire

1. Have you ever taken a college-level physics course? ____ If so, please list the courses that you have taken:

____

____

____

____

____

2. Have you ever taken a physics course in high school? ____ If so, please list and indicate if it was an advance placement (AP) or college preparatory (CP) course:

____

____

3. Have you ever taken a college-level mathematics or statistics course? ____ If so, please list the courses that you have taken:

____

____

____

____

____

4. Have you ever taken a mathematics or statistics course in high school? ____ If so, please list and indicate if it was an advance placement (AP) or college preparatory (CP) course:

____

---

5. Do you experience math anxiety?

   a. yes, frequently

   b. occasionally

   c. rarely

   d. no, not at all

   e. not sure

APPENDIX D

Old, Near, and Far Transfer Problems

Old problem, kinematics

A cat jumps from the roof of a building, uniformly accelerating downward at a rate of 10 m/s^2, with an average velocity of 5 m/s. If the cat is travelling 2 m/s when it passes one window and 8 m/s when it passes another window, how far apart are the windows?

Near transfer problem, kinematics

A train increases its velocity from 10 m/s to 30 m/s by uniformly accelerating at the rate of 4 m/s^2. During this period of time the train has an average velocity of 20 m/s. How far did the train travel during this same period of time?

Far transfer problem, kinematics

A rocket fires its booster, causing it to uniformly accelerate 20 m/s^2 over a distance of 560 m. During this time, the rocket's average velocity was 140 m/s. What was the rocket's change in velocity during this time?

Old problem, dynamics

A weightlifter performs 1950 Nm of work when he lifts a 130 kg barbell a distance of 1.5 m. What is the acceleration of the barbell?

Near transfer problem, dynamics

A 1000 kg car is pushed a distance of 5 m. If 2000 Nm of work is required to push the car, what is its acceleration?

Far transfer problem, dynamics

A 50 kg piston moves .1 m with a uniform acceleration of 2 m/s^2. How much work can the piston perform?

APPENDIX E

Table E1. Students' PF similarities among categories.

STUDENT	Overall PF Similarity	Kinematics PF Similarity	Dynamics PF Similarity	Kinematics Equation 1 PF Similarity	Kinematics Equation 2 PF Similarity	Dynamics Equation 1 PF Similarity	Dynamics Equation 2 PF Similarity
14	0.20	0.50	0.20	1.00	0	0.50	0
15	0.26	0.33	0.50	0.33	0.33	0.67	0.33
16	0.21	0.25	0.33	0	0.50	0.33	0.33
17	0.24	0.40	0.33	0.33	0.50	0.33	0.33
18	0.21	0.67	0.20	0.50	1.00	0.33	0
19	0.28	0.50	0.60	1.00	0	0.67	0.50
20	0.13	0.20	0.17	0.50	0	0.33	0
21	0.23	0.20	0.50	0.33	0	0.67	0.33
22	0.17	0.40	0.25	0.33	0.50	0.50	0
23	0.19	0	0.60	0	0	1.00	0.33
24	0.38	0.20	0.60	0	0.50	0.50	0.67
25	0.31	0.25	0.80	0.33	0	1.00	0.67
26	0.29	0.40	0.50	0.33	0.50	0.67	0.33
27	0.29	0.50	0.33	0.67	0	0.33	0.33
28	0.23	0.33	0.20	0	1.00	0.50	0
29	0.24	0.40	0.40	0.33	0.50	0.33	0.50
30	0.24	0.75	0.33	1.00	0.50	0.33	0.33

31	0.24	0.20	0.50	0	0.50	0.67	0.33
32	0.18	0.20	0.60	0.33	0	0.67	0.50
33	0.07	0	0.25	0	0	0.50	0
34	0.35	0.50	0.67	0.50	0.50	0.67	0.67
35	0.28	0.75	0.33	1.00	0.50	0.33	0.33
36	0.29	0.75	0.33	1.00	0.50	0.33	0.33
37	0.21	0.33	0.40	0.50	0	0.50	0.33
38	0.13	0	0.40	0	0	0.33	0.50
39	0.35	0.40	1.00	0.33	0.50	1.00	1.00
40	0.22	0.40	0.40	0.33	0.50	0.67	0
41	0.24	0.40	0.33	0.33	0.50	0.33	0.33
42	0.20	0.25	0.50	0	0.50	0.67	0.33
45	0.35	0.75	0.50	1.00	0.50	0.67	0.33
51	0.29	0.75	0.33	1.00	0.50	0.33	0.33
54	0.27	0.40	0.67	0.33	0.50	0.67	0.67
61	0.15	0.40	0.17	1.00	0	0	0.33
62	0.33	0.50	0.33	0.50	0.50	0.33	0.33
65	0.32	0.40	0.60	0.50	0.33	1.00	0.33
66	0.19	0.20	0.60	0.33	0	0.67	0.50
68	0.33	0.67	0.40	0.50	1.00	0.33	0.50
69	0	0	0	0	0	0	0
71	0.40	0.50	0.60	0.50	0.50	1.00	0.33
74	0.18	0.20	0.50	0.33	0	0.67	0.33
80	0.13	0.20	0.20	0	0.50	0.33	0
81	0.27	0.50	0.33	1.00	0	0.33	0.33
82	0.19	0.50	0	0.50	0.50	0	0
83	0.22	0.50	0.50	0.50	0.50	0.50	0.50
84	0.21	0.25	0.33	0.50	0	0.33	0.33
85	0.29	0.50	0.67	0.67	0.33	0.67	0.67
86	0.20	0	0.50	0	0	0	1.00
Mean	0.2379	0.3762	0.4209	0.4353	0.3296	0.4998	0.3498

Table E2. Descriptive statistics for PFSim categories.

	N	Mean		Std. Deviation	Variance
	Statistic	Statistic	Std. Error	Statistic	Statistic
Overall PFsim	47	.24	.01	.08	.01

Kinematics PFsim	47	.38	.03	.21	.05
Dynamics PFsim	47	.42	.03	.20	.04
Kinematics Eqn.1 PFsim	47	.44	.05	.34	.11
Kinematics Eqn.2 PFsim	47	.33	.04	.29	.09
Dynamics Eqn.1 PFsim	47	.50	.04	.26	.07
Dynamics Eqn.2 PFsim	47	.35	.04	.24	.06

Table E3. Problem Solving Scores

STUDENT #	Overall Problem Scores	Kinematics Scores	Dynamics Scores	Kinematics Equation1 Scores	Kinematics Equation2 Scores	Dynamics Equation1 Scores	Dynamics Equation2 Scores
14	34	16	18	8	8	10	8
15	37	24	13	12	12	5	8
16	8	8	0	4	4	0	0
17	28	12	16	6	6	8	9
18	24	8	16	4	4	8	8
19	39	16	23	8	8	13	11
20	24	8	16	4	4	9	8
21	17	4	13	0	4	8	5
22	42	24	18	12	12	12	6
23	23	3	20	1	2	13	8
24	38	18	20	9	9	11	10
25	48	24	24	12	12	13	12
26	28	9	19	5	4	11	9
27	33	9	24	4	5	13	12
28	19	13	6	8	5	5	1
29	21	5	16	4	1	8	8
30	26	2	24	1	1	12	12
31	30	15	15	3	12	3	12
32	29	9	20	9	0	12	8
33	22	11	11	6	5	3	9
34	30	14	16	8	6	7	9

35	44	20	24	8	12	12	12
36	30	14	16	6	8	8	8
37	32	14	18	7	7	10	8
38	24	8	16	4	4	8	8
39	35	16	19	6	10	11	9
40	25	9	16	5	4	8	8
41	36	20	16	12	8	9	8
42	12	4	8	4	0	4	4
45	41	17	24	8	9	12	12
51	37	20	17	10	10	8	9
54	25	5	20	2	3	11	10
61	30	6	24	0	6	12	12
62	29	11	18	3	8	9	10
65	40	16	24	8	8	12	12
66	34	20	14	10	10	8	7
68	48	24	24	12	12	13	12
69	19	11	8	6	5	4	4
71	39	19	20	9	10	10	10
74	22	9	13	3	6	6	7
80	26	11	15	5	6	8	8
81	32	16	16	8	8	8	8
82	32	12	20	8	4	12	8
83	12	4	8	4	0	4	4
84	33	9	24	5	4	13	12
85	33	16	17	8	8	9	8
86	24	12	12	8	4	8	4

Table E4. Problem Solving Scores Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Overall.Problem.Solving	47	8.00	48.00	29.6596	8.96437
Kinematics.Scores	47	2.00	24.00	12.6596	6.02270
Dynamics.Scores	47	.00	24.00	17.0000	5.38113
Kinematics.Equation1.Scores	47	.00	12.00	6.3191	3.23101
Kinematics.Equation2.Scores	47	.00	12.00	6.3404	3.46584
Dynamics.Equation1.Scores	47	.00	12.00	8.8298	3.05258
Dynamics.Equation2.Scores	47	.00	12.00	8.4043	2.85642
Valid N (listwise)	47				

Table E5. Pearson Correlations between Problem Solving Performance and PFsims.

		Overall.Problem. Performance	Overall.PF.Similarity
Pearson Correlation	Overall.Problem.Performance	1.000	.525
	Overall.PF.Similarity	.525	1.000
Sig. (1-tailed)	Overall.Problem.Performance	.	.000
	Overall.PF.Similarity	.000	.
		Far.Transfer. Performance	Overall.PF.Similarity
Pearson Correlation	Far.Transfer.Performance	1.000	.550
	Overall.PF.Similarity	.550	1.000
Sig. (1-tailed)	Far.Transfer.Performance	.	.000
	Overall.PF.Similarity	.000	.
		Near.Transfer. Performance	Overall.PF.Similarity
Pearson Correlation	Near.Transfer.Performance	1.000	.275
	Overall.PF.Similarity	.275	1.000
Sig. (1-tailed)	Near.Transfer.Performance	.	.030
	Overall.PF.Similarity	.030	.
		Old.Transfer. Performance	Overall.PF.Similarity
Pearson Correlation	Old.Transfer.Performance	1.000	.061
	Overall.PF.Similarity	.061	1.000
Sig. (1-tailed)	Old.Transfer.Performance	.	.342

	Overall.PF.Similarity	.342	.
		Far.Kinematics	Kinematics.PF. Similarity
Pearson Correlation	Far.Kinematics	1.000	.217
	Kinematics.PF.Similarity	.217	1.000
Sig. (1-tailed)	Far.Kinematics	.	.072
	Kinematics.PF.Similarity	.072	.
		Kinematics.PF. Similarity	Old.Kinematics
Pearson Correlation	Kinematics.PF.Similarity	1.000	.072
	Old.Kinematics	.072	1.000
Sig. (1-tailed)	Kinematics.PF.Similarity	.	.315
	Old.Kinematics	.315	.
		Kinematics.PF. Similarity	Near.Kinematics
Pearson Correlation	Kinematics.PF.Similarity	1.000	.135
	Near.Kinematics	.135	1.000
Sig. (1-tailed)	Kinematics.PF.Similarity	.	.182
	Near.Kinematics	.182	.
		Far.Dynamics	Dynamics.PF. Similarity
Pearson Correlation	Far.Dynamics	1.000	.197
	Dynamics.PF.Similarity	.197	1.000
Sig. (1-tailed)	Far.Dynamics	.	.092
	Dynamics.PF.Similarity	.092	.
		Dynamics.PF. Similarity	Old.Dynamics

Pearson Correlation	Dynamics.PF.Similarity	1.000	.117
	Old.Dynamics	.117	1.000
Sig. (1-tailed)	Dynamics.PF.Similarity	.	.217
	Old.Dynamics	.217	.
		Dynamics.PF. Similarity	Near.Dynamics
Pearson Correlation	Dynamics.PF.Similarity	1.000	.043
	Near.Dynamics	.043	1.000
Sig. (1-tailed)	Dynamics.PF.Similarity	.	.386
	Near.Dynamics	.386	.
		Far.step1	Kinematics1.PFsim
Pearson Correlation	Far.step1	1.000	.068
	Kinematics1.PFsim	.068	1.000
Sig. (1-tailed)	Far.step1	.	.324
	Kinematics1.PFsim	.324	.
		Far.step2	Kinematics2.PFsim
Pearson Correlation	Far.step2	1.000	.027
	Kinematics2.PFsim	.027	1.000
Sig. (1-tailed)	Far.step2	.	.428
	Kinematics2.PFsim	.428	.
		Far.Dynamics.step1	Dynamics1.PFsim
Pearson Correlation	Far.Dynamics.step1	1.000	.105
	Dynamics1.PFsim	.105	1.000
Sig. (1-tailed)	Far.Dynamics.step1	.	.242
	Dynamics1.PFsim	.242	.

		Far.Dynamics.step2	Dynamics2.PFsim
Pearson Correlation	Far.Dynamics.step2	1.000	.207
	Dynamics2.PFsim	.207	1.000
Sig. (1-tailed)	Far.Dynamics.step2	.	.081
	Dynamics2.PFsim	.081	.

Table E6. Students problem solving scores within different categories.

STUDENT	Old Transfer Kinematics Problem							
	Choose Equation 1	Manipulate Equation 1	Substitute Equation 1	Calculate Equation 1	Choose Equation 2	Manipulate Equation 2	Substitute Equation 2	Calculate Equation 2
14	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0
18	1	1	1	1	1	1	1	1
19	0	0	0	0	0	0	0	0
20	1	1	1	1	1	1	1	1
21	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0
25	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1
28	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0

30	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0
33	1	1	1	1	1	1	1	1
34	0	0	0	0	0	0	0	0
35	1	1	1	1	1	1	1	1
36	0	0	0	0	0	0	0	0
37	1	1	1	1	1	1	1	1
38	0	0	0	0	0	0	0	0
39	1	1	1	1	1	1	1	1
40	0	0	0	0	0	0	0	0
41	1	1	1	1	0	0	0	0
42	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0
51	1	1	1	1	1	1	1	1
54	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0
62	0	0	1	0	1	1	1	1
65	0	0	0	0	0	0	0	0
66	1	1	1	1	1	1	1	1
68	1	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1	1
71	0	0	0	0	0	0	0	0
74	0	0	0	0	0	0	0	0
80	0	0	0	0	0	1	0	0
81	0	0	0	0	0	0	0	0

82	1	1	1	1	1	1	1	1
83	0	0	0	0	0	0	0	0
84	1	1	1	1	1	1	1	1
85	0	0	0	0	0	0	0	0
86	1	1	1	1	1	1	1	1

STUDENT	Old Transfer Dynamics Problem							
	Choose Equation 1	Manipulate Equation 1	Substitute Equation 1	Calculate Equation 1	Choose Equation 2	Manipulate Equation 2	Substitute Equation 2	Calculate Equation 2
14	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1
21	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0
23	1	1	1	1	1	1	1	1
24	0	0	0	0	0	0	0	0
25	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1
28	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0

30	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0
33	1	1	1	1	1	1	1	1
34	0	0	0	0	0	0	0	0
35	1	1	1	1	1	1	1	1
36	0	0	0	0	0	0	0	0
37	1	1	1	1	1	1	1	1
38	0	0	0	0	0	0	0	0
39	0	0	1	0	1	1	1	1
40	0	0	0	0	0	0	0	0
41	1	1	1	1	1	1	1	1
42	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0
51	1	1	1	1	1	1	1	1
54	1	1	1	1	1	1	1	1
61	0	0	0	0	0	0	0	0
62	0	0	0	0	1	1	0	0
65	0	0	0	0	0	0	0	0
66	1	1	1	1	1	1	1	1
68	1	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1	1
71	0	0	0	0	0	0	0	0
74	0	0	0	0	0	0	0	0
80	1	1	1	1	1	1	1	1
81	0	0	0	0	0	0	0	0

82	1	1	1	1	0	0	0	0
83	0	0	0	0	0	0	0	0
84	1	1	1	1	1	1	1	1
85	0	0	0	0	0	0	0	0
86	1	1	1	1	0	0	0	0

STUDENT	Near Transfer Kinematics Problem							
	Choose Equation 1	Manipulate Equation 1	Substitute Equation 1	Calculate Equation 1	Choose Equation 2	Manipulate Equation 2	Substitute Equation 2	Calculate Equation 2
14	1	1	1	1	1	1	1	1
15	2	2	2	2	2	2	2	2
16	1	1	1	1	1	1	1	1
17	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1
20	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0
22	2	2	2	2	2	2	2	2
23	0	0	0	0	2	0	0	0
24	0	0	0	0	0	0	0	0
25	1	1	1	1	1	1	1	1
26	1	0	0	0	0	0	0	0
27	0	0	0	0	1	0	0	0
28	1	1	1	1	1	1	1	1
29	0	0	0	0	0	0	0	1

30	1	0	0	0	1	0	0	0
31	0	2	0	0	2	2	2	2
32	2	1	1	1	0	0	0	0
33	1	1	1	1	1	1	1	1
34	1	0	0	0	0	0	0	0
35	1	1	1	1	1	1	1	1
36	0	0	0	0	0	0	0	0
37	1	1	1	1	1	1	1	1
38	1	1	1	1	1	1	1	1
39	0	0	0	0	0	0	0	0
40	0	1	0	0	0	0	0	0
41	1	1	1	1	1	1	1	1
42	1	1	1	1	0	0	0	0
45	2	2	2	2	2	2	2	2
51	1	1	1	1	1	1	1	1
54	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	2
62	0	1	1	0	1	1	1	1
65	1	1	1	1	1	1	1	1
66	1	1	1	1	1	1	1	1
68	1	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1	1
71	1	0	2	2	2	0	2	2
74	0	0	0	0	0	0	0	0
80	1	1	1	1	1	1	1	1

81	0	0	2	2	0	0	2	2
82	0	0	0	0	0	0	0	0
83	1	1	1	1	0	0	0	0
84	0	0	0	0	0	0	0	0
85	0	1	1	2	0	0	2	2
86	0	0	0	0	0	0	0	0

STUDENT	Near Transfer Dynamics Problem							
	Choose Equation 1	Manipulate Equation 1	Substitute Equation 1	Calculate Equation 1	Choose Equation 2	Manipulate Equation 2	Substitute Equation 2	Calculate Equation 2
14	2	2	2	2	2	0	2	2
15	1	1	1	1	0	1	1	1
16	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0
18	1	1	1	1	1	1	1	1
19	1	1	1	1	0	1	1	1
20	0	0	0	0	0	0	0	0
21	1	1	1	1	1	1	1	1
22	2	2	2	2	2	0	0	0
23	1	1	1	1	1	1	1	1
24	0	0	0	0	0	0	0	0
25	1	1	1	1	1	1	1	1
26	0	0	0	0	0	0	0	0
27	1	1	1	1	1	1	1	1
28	1	1	1	1	1	0	0	0
29	1	1	1	1	0	2	1	1

30	2	2	2	2	2	2	2	2
31	2	0	0	0	2	2	2	2
32	2	2	2	2	1	1	1	1
33	0	0	0	0	1	1	1	1
34	1	1	1	1	1	1	1	1
35	1	1	1	1	1	1	1	1
36	1	1	1	1	1	1	1	1
37	1	1	1	1	1	1	1	1
38	2	2	2	2	2	2	2	2
39	1	1	1	1	0	0	0	0
40	1	1	1	1	1	1	1	1
41	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0
45	2	2	2	2	2	2	2	2
51	1	1	1	1	1	1	1	1
54	1	0	0	0	1	0	0	0
61	2	2	2	2	2	2	2	2
62	1	1	1	1	1	1	1	1
65	2	2	2	2	2	2	2	2
66	1	1	1	1	1	1	1	1
68	1	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1	1
71	1	1	2	2	1	1	2	2
74	0	0	0	0	0	0	0	0
80	1	1	1	1	1	1	1	1

81	0	0	2	2	0	0	2	2
82	1	1	1	1	1	1	1	1
83	0	0	0	0	0	0	0	0
84	1	1	1	1	1	1	1	1
85	0	1	2	2	0	0	2	2
86	0	0	0	0	0	0	0	0

STUDENT	Far Transfer Kinematics Problem							
	Choose Equation 1	Manipulate Equation 1	Substitute Equation 1	Calculate Equation 1	Choose Equation 2	Manipulate Equation 2	Substitute Equation 2	Calculate Equation 2
14	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1
16	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1
20	0	0	0	0	0	0	0	0
21	0	0	0	0	1	1	1	1
22	1	1	1	1	1	1	1	1
23	0	1	0	0	0	0	0	0
24	0	1	0	0	0	0	0	0
25	1	1	1	1	1	1	1	1
26	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0
28	1	1	1	1	1	0	0	0
29	1	1	1	1	0	0	0	0

30	0	0	0	0	0	0	0	0
31	0	1	0	0	1	1	1	1
32	1	1	1	1	0	0	0	0
33	0	0	0	0	0	0	0	0
34	1	0	0	0	0	0	0	0
35	0	0	0	0	1	1	1	1
36	0	0	0	0	0	0	0	0
37	1	1	1	1	1	1	1	1
38	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0
40	1	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1	1
42	0	0	0	0	0	0	0	0
45	0	0	0	0	1	0	0	0
51	1	1	1	1	1	1	1	1
54	0	1	0	0	0	0	0	0
61	0	0	0	0	1	1	1	1
62	0	0	0	0	0	0	0	0
65	1	1	1	1	1	1	1	1
66	1	1	1	1	1	1	1	1
68	1	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1	1
71	1	1	1	1	1	1	1	1
74	0	0	0	0	0	0	0	0
80	1	1	1	1	1	1	1	1

81	1	1	1	1	1	1	1	1
82	1	1	1	1	0	0	0	0
83	0	0	0	0	0	0	0	0
84	1	0	0	0	0	0	0	0
85	1	1	1	1	1	1	1	1
86	1	1	1	1	0	0	0	0

STUDENT	Far Transfer Dynamics Problem							
	Choose Equation 1	Manipulate Equation 1	Substitute Equation 1	Calculate Equation 1	Choose Equation 2	Manipulate Equation 2	Substitute Equation 2	Calculate Equation 2
14	0	0	1	1	0	0	1	1
15	0	0	0	1	2	1	1	1
16	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0
18	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1
21	1	1	1	1	0	1	0	0
22	1	1	1	1	1	1	1	1
23	1	1	1	1	0	0	0	0
24	1	0	1	0	1	1	1	1
25	1	1	1	1	1	1	1	1
26	2	0	2	2	1	0	2	2
27	1	1	1	1	1	1	1	1
28	1	0	0	0	0	0	0	0

29	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1
31	1	0	0	0	1	1	1	1
32	1	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1	1
34	0	0	0	0	0	0	0	0
35	1	1	1	1	1	1	1	1
36	1	1	1	1	1	1	1	1
37	1	1	1	1	1	1	1	1
38	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0
40	1	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1	1
42	1	1	1	1	1	1	1	1
45	1	1	1	1	1	1	1	1
51	1	1	1	1	1	1	1	1
54	1	0	0	0	1	0	0	0
61	1	1	1	1	1	1	1	1
62	1	1	1	1	1	1	1	1
65	1	1	1	1	1	1	1	1
66	1	1	1	1	1	1	1	1
68	1	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1	1
71	1	1	1	1	1	1	1	1
74	0	0	0	0	1	1	1	1

80	1	1	1	1	1	1	1	1
81	1	1	1	1	1	1	1	1
82	1	1	1	1	1	1	1	1
83	1	1	1	1	1	1	1	1
84	1	1	1	1	1	1	1	1
85	1	1	1	1	1	1	1	1
86	1	1	1	1	1	1	1	1

STUDENT	**OVERALL LEVEL**				
	Overall PF Similarity	Overall Problem Performance	Old Transfer Performance	Near Transfer Performance	Far Transfer Performance
14	0.20	34	0	22	12
15	0.26	37	0	23	14
16	0.21	8	0	8	0
17	0.24	0	0	0	0
18	0.21	24	8	8	8
19	0.28	39	8	15	16
20	0.13	24	16	0	8
21	0.23	17	0	8	9
22	0.17	42	0	26	16
23	0.19	23	8	10	5
24	0.38	7	0	0	7
25	0.31	48	16	16	16
26	0.29	28	16	1	11
27	0.29	33	16	9	8

28	0.23	19	0	13	6
29	0.24	21	0	9	12
30	0.24	26	0	18	8
31	0.24	30	0	20	10
32	0.18	29	0	17	12
33	0.07	36	16	12	8
34	0.35	10	0	9	1
35	0.28	44	16	16	12
36	0.29	16	0	8	8
37	0.21	48	16	16	16
38	0.13	24	0	24	0
39	0.35	17	13	4	0
40	0.22	25	0	9	16
41	0.24	36	12	8	16
42	0.20	12	0	4	8
45	0.35	41	0	32	9
51	0.29	48	16	16	16
54	0.27	13	8	2	3
61	0.15	30	0	18	12
62	0.33	29	7	14	8
65	0.32	40	0	24	16
66	0.19	48	16	16	16
68	0.33	48	16	16	16
69	-	48	16	16	16

71	0.40	39	0	23	16
74	0.18	4	0	0	4
80	0.13	41	9	16	16
81	0.27	32	0	16	16
82	0.19	32	12	8	12
83	0.22	12	0	4	8
84	0.21	33	16	8	9
85	0.29	33	0	17	16
86	0.20	24	12	0	12

STUDENT	**DOMAIN LEVEL**							
	Kinematics PF Similarity	Old Transfer Performance	Near Transfer Performance	Far Transfer Performance	Dynamics PF Similarity	Old Transfer Performance	Near Transfer Performance	Far Transfer Performance
14	0.50	0	8	8	0.20	0	14	4
15	0.33	0	16	8	0.50	0	7	6
16	0.25	0	8	0	0.33	0	0	0
17	0.40	0	0	0	0.33	0	0	0
18	0.67	8	0	0	0.20	0	8	8
19	0.50	0	8	8	0.60	8	7	8
20	0.20	8	0	0	0.17	8	0	8
21	0.20	0	0	4	0.50	0	8	5
22	0.40	0	16	8	0.25	0	10	8
23	-	0	2	1	0.60	8	8	4
24	0.20	0	0	1	0.60	0	0	6

25	0.25	8	8	8	0.80	8	8	8
26	0.40	8	1	0	0.50	8	0	11
27	0.50	8	1	0	0.33	8	8	8
28	0.33	0	8	5	0.20	0	5	1
29	0.40	0	1	4	0.40	0	8	8
30	0.75	0	2	0	0.33	0	16	8
31	0.20	0	10	5	0.50	0	10	5
32	0.20	0	5	4	0.60	0	12	8
33	-	8	8	0	0.25	8	4	8
34	0.50	0	1	1	0.67	0	8	0
35	0.75	8	8	4	0.33	8	8	8
36	0.75	0	0	0	0.33	0	8	8
37	0.33	8	8	8	0.40	8	8	8
38	-	0	8	0	0.40	0	16	0
39	0.40	8	0	0	1.00	5	4	0
40	0.40	0	1	8	0.40	0	8	8
41	0.40	4	8	8	0.33	8	0	8
42	0.25	0	4	0	0.50	0	0	8
45	0.75	0	16	1	0.50	0	16	8
51	0.75	8	8	8	0.33	8	8	8
54	0.40	0	0	1	0.67	8	2	2
61	0.40	0	2	4	0.17	0	16	8
62	0.50	5	6	0	0.33	2	8	8
65	0.40	0	8	8	0.60	0	16	8

66	0.20	8	8	8	0.60	8	8	8
68	0.67	8	8	8	0.40	8	8	8
69	-	8	8	8	-	8	8	8
71	0.50	0	11	8	0.60	0	12	8
74	0.20	0	0	0	0.50	0	0	4
80	0.20	1	8	8	0.20	8	8	8
81	0.50	0	8	8	0.33	0	8	8
82	0.50	8	0	4	-	4	8	8
83	0.50	0	4	0	0.50	0	0	8
84	0.25	8	0	1	0.33	8	8	8
85	0.50	0	8	8	0.67	0	9	8
86	-	8	0	4	0.50	4	0	8

STUDENT	**EQUATION LEVEL**							
	Dynamics Equation 1 PF Similarity	Old Transfer Step 1 Perfomance	Near Transfer Step 1 Perfomance	Far Transfer Step 1 Perfomance	Dynamics Equation 2 PF Similarity	Old Transfer Step 2 Perfomance	Near Transfer Step 2 Perfomance	Far Transfer Step 2 Perfomance
14	0.50	0	8	2	-	0	6	2
15	0.67	0	4	1	0.33	0	3	5
16	0.33	0	0	0	0.33	0	0	0
17	0.33	0	0	0	0.33	0	0	0
18	0.33	0	4	4	-	0	4	4
19	0.67	5	4	4	0.50	4	3	4
20	0.33	5	0	4	-	4	0	4
21	0.67	0	4	4	0.33	0	4	1

22	0.50	0	8	4	-	0	2	4
23	1.00	5	4	4	0.33	4	4	0
24	0.50	0	0	2	0.67	0	0	4
25	1.00	5	4	4	0.67	4	4	4
26	0.67	5	0	6	0.33	4	0	5
27	0.33	5	4	4	0.33	4	4	4
28	0.50	0	4	1	-	0	1	0
29	0.33	0	4	4	0.50	0	4	4
30	0.33	0	8	4	0.33	0	8	4
31	0.67	0	2	1	0.33	0	8	4
32	0.67	0	8	4	0.50	0	4	4
33	0.50	5	0	4	-	4	4	4
34	0.67	0	4	0	0.67	0	4	0
35	0.33	5	4	4	0.33	4	4	4
36	0.33	0	4	4	0.33	0	4	4
37	0.50	5	4	4	0.33	4	4	4
38	0.33	0	8	0	0.50	0	8	0
39	1.00	2	4	0	1.00	4	0	0
40	0.67	0	4	4	-	0	4	4
41	0.33	5	0	4	0.33	4	0	4
42	0.67	0	0	4	0.33	0	0	4
45	0.67	0	8	4	0.33	0	8	4
51	0.33	5	4	4	0.33	4	4	4
54	0.67	5	1	1	0.67	4	1	1

61	-	0	8	4	0.33	0	8	4
62	0.33	1	4	4	0.33	2	4	4
65	1.00	0	8	4	0.33	0	8	4
66	0.67	5	4	4	0.50	4	4	4
68	0.33	5	4	4	0.50	4	4	4
69	-	5	4	4	-	4	4	4
71	1.00	0	6	4	0.33	0	6	4
74	0.67	0	0	0	0.33	0	0	4
80	0.33	5	4	4	-	4	4	4
81	0.33	0	4	4	0.33	0	4	4
82	-	4	4	4	-	0	4	4
83	0.50	0	0	4	0.50	0	0	4
84	0.33	5	4	4	0.33	4	4	4
85	0.67	0	5	4	0.67	0	4	4
86	-	4	0	4	1.00	0	0	4

STUDENT	EQUATION LEVEL							
	Kinematics Equation 1 PF Similarity	Old Transfer Step 1 Performance	Near Transfer Step 1 Performance	Far Transfer Step 1 Performance	Kinematics Equation 2 PF Similarity	Old Transfer Step 2 Performance	Near Transfer Step 2 Performance	Far Transfer Step 2 Performance
14	1.00	0	4	4	-	0	4	4
15	0.33	0	8	4	0.33	0	8	4
16	-	0	4	0	0.50	0	4	0
17	0.33	0	0	0	0.50	0	0	0
18	0.50	4	0	0	1.00	4	0	0
19	1.00	0	4	4	-	0	4	4

20	0.50	4	0	0	-	4	0	0
21	0.33	0	0	0	-	0	0	4
22	0.33	0	8	4	0.50	0	8	4
23	-	0	0	1	-	0	2	0
24	-	0	0	1	0.50	0	0	0
25	0.33	4	4	4	-	4	4	4
26	0.33	4	1	0	0.50	4	0	0
27	0.67	4	0	0	-	4	1	0
28	-	0	4	4	1.00	0	4	1
29	0.33	0	0	4	0.50	0	1	0
30	1.00	0	1	0	0.50	0	1	0
31	-	0	2	1	0.50	0	8	4
32	0.33	0	5	4	-	0	0	0
33	-	4	4	0	-	4	4	0
34	0.50	0	1	1	0.50	0	0	0
35	1.00	4	4	0	0.50	4	4	4
36	1.00	0	0	0	0.50	0	0	0
37	0.50	4	4	4	-	4	4	4
38	-	0	4	0	-	0	4	0
39	0.33	4	0	0	0.50	4	0	0
40	0.33	0	1	4	0.50	0	0	4
41	0.33	4	4	4	0.50	0	4	4
42	-	0	4	0	0.50	0	0	0
45	1.00	0	8	0	0.50	0	8	1

51	1.00	4	4	4	0.50	4	4	4
54	0.33	0	0	1	0.50	0	0	0
61	1.00	0	0	0	-	0	2	4
62	0.50	1	2	0	0.50	4	4	0
65	0.50	0	4	4	0.33	0	4	4
66	0.33	4	4	4	-	4	4	4
68	0.50	4	4	4	1.00	4	4	4
69	-	4	4	4	-	4	4	4
71	0.50	0	5	4	0.50	0	6	4
74	0.33	0	0	0	-	0	0	0
80	-	0	4	4	0.50	1	4	4
81	1.00	0	4	4	-	0	4	4
82	0.50	4	0	4	0.50	4	0	0
83	0.50	0	4	0	0.50	0	0	0
84	0.50	4	0	1	-	4	0	0
85	0.67	0	4	4	0.33	0	4	4
86	-	4	0	4	-	4	0	0