Models of Learning Progress in Solving Complex Problems: Expertise Development in Teaching and Learning

Abstract

This study proposes that learning is a process of transitioning from one stage to another stage within a knowledge base that features concepts and relations. Drawing on the theories of expertise, this study explored two different models of learning progress (i.e., three- and two-stage models) in the context of classroom learning and identified a model that was a good fit to the data. Participants in this investigation included 136 students and 7 professors from six different universities in the United States. In order to detect and validate stages of learning progress in participants' written responses to an ill-structured and complex problem scenario, this study utilized Exploratory Factor Analysis (EFA) and the Continuous Log-Linear Cognitive Diagnostic Model (C-LCDM) method (Bozard, 2010). The results demonstrate that the three latent classes matched the three stages of the three-stage model. This study provides an account of a diagnostic model of learning progress and associated assessment methods, yet further studies are required to investigate different conditions.

Keywords: Learning progress, cognitive change, problem solving, mental models, latent class

Models of Learning Progress in Solving Complex Problems:

Expertise Development in Teaching and Learning

1. Introduction

Learning is fundamentally about the change in knowledge and skills needed to solve problems (Bransford et al., 2000; Spector, 2004). Over the past few decades, many studies have addressed how people learn through complex problem solving in diverse disciplines, including cognitive psychologies (Sinnott, 1989), applied psychologies (Zsambok & Klein, 1997), and educational psychologies (Jonassen, 1997; Spiro et al., 1996), yet understanding changes in the ability to solve a problem within the context of classroom teaching remains a challenge due to the uncertain and complex nature of problems (Choi & Lee, 2009) and changes that take place in the short term.

The theories of mental models and expertise development can help address these issues. The theory of mental models explains that problem solving involves a process of building mental representations of a problem situation (Johnson-Laird, 1983). Mental models are holistic structural representations of the facts, concepts, variables, objects, and their relationships within a problem situation (Dochy et al., 2003; Jonassen et al., 1993; Segers, 1997). Cognitive change takes place when learners confront unfamiliar and challenging situations (diSessa, 2006; Festinger, 1962; Piaget, 1964) or a pre-structural lack of knowledge (Biggs & Collis, 1982). When striving to resolve problem situations, learners experience changes in their mental representations whereby the problem situations are recognized, defined, and organized (Seel, 2003, 2004).

Learners possibly experience qualitatively different levels of knowledge structure when engaged in problem solving. In line with the idea that children experience qualitatively distinct but sequential knowledge states (Piaget, 1964), developmental psychologists have seen that learning and development evolve as the learner constructs a qualitatively distinct knowledge structure (Alexander, 2003, 2004; Flavell & Miller, 1998; Siegler, 2005; Siegler et al., 2009; Werner, 1957; Vygotsky, 1934/1978). A number of experimental studies have demonstrated that qualitatively different cognitive stages take place when learners respond to problems in both the short term and the long term (e.g., Alexander & Murphy, 1998; Chen & Siegler, 2000; Opfer & Siegler, 2004; Siegler et al., 2009; Vosniadou et al., 2008).

Expertise studies have sought evidence to explain the development of expertise. As Figure 1 illustrates, traditional expert-novice studies define an expert as one who consistently and successfully performs in a specific, selected domain due to a highly enhanced, efficient, and effective long-term memory (Ericsson & Kintsch, 1995; Flavell, 1985; Simon & Chase, 1973). In addition, contemporary studies of expertise do not restrict expertise development to chunking and patter recognition. For example, Ericsson (2003, 2005, 2006) suggested that expertise is developed by deliberate practice in which learners engage in appropriate and challenging tasks carefully chosen by masters, devote years of practice to improve their performance, and refine their cognitive mechanisms through self-monitoring efforts, such as planning, reflection, and evaluation. From this point of view, expertise development in a particular domain requires long-term devotion (e.g., ten or more years) to disciplined and focused practice (Ericsson, 2003, 2005, 2006).

This study addresses two major limitations often found in the traditional approach. One is that traditional studies tend to contrast two extremes, experts and novices, resulting in “lack of developmental focus" (Alexander, 2004, p. 278). Missing accounts of the middle stages leave the developmental process somewhat unclear. Furthermore, Alexander (2004) argued that the educators cannot confidently apply the findings of research conducted in a non-academic setting.

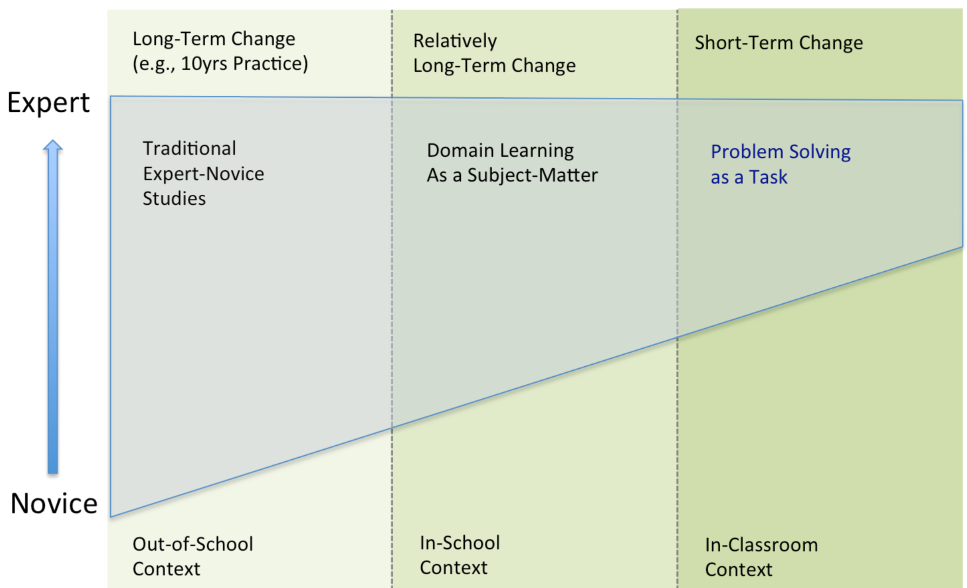

Figure 1. The focus of expertise development studies.

In response to these problems, Alexander (2003, 2004) developed a Model of Domain Learning (MDL) to explain multiple stages of expertise development in the educational context and validated the model (Alexander & Murphy, 1998; Alexander et al., 2004). In spite of the unique value of the model, it still assumes that long-term changes during schooling are required to master a domain.

The current study shifts the focus of expertise development to problem-solving situations (i.e., task level) in the classroom and examines how short-term changes can lead to expertise. Several studies (e.g., Chi, 2006; Newell & Simon, 1972) have suggested that an individual's understanding of a problem situation reflects levels of expertise in solving that problem. Based on the same idea, Gobet and Wood (1999) tested an explicit knowledge model of learning and expertise for the design of computer-based tutoring systems. Concerning problem solving in the classroom, recent studies have investigated differences between and changes in expertise (Ifenthaler et al., 2009; Schlomske & Pirnay-Dummer, 2008; Spector & Koszalka, 2004). For example, Ifenthaler et al. (2009) investigated longitudinal changes using hierarchical linear modeling (HLM) techniques. However, a common limitation of these studies is a "lack of developmental focus" (Alexander, 2004, p. 278), either simplifying the difference to "expert vs. novice" or falling short of explicitly modeling stages of expertise development.

Using confirmatory analysis, the current study explored potential models of learning progress in solving ill-structured, complex problem situations in a classroom setting. A valid model can provide better insight into the cognitive development associated with solving complex problems, insight that is essential to adapt teaching and learning to individual differences in learning with precision and confidence (Grow-Maienza et al., 2001; Hattie, 2009; Stigler & Stevenson, 1991). The following research questions guided the study:

1. Which theories explain learning progress as expertise development in solving a complex problem?
2. What are some potential models of learning progress that are both theoretically rigorous and viable?
3. Which model best explains the stages of learning progress?

2. Learning Progress in Solving Ill-Structured and Complex Problems

2.1 Ill-Structured and Complex Problems

Problems in professional and daily life are often characterized by their uncertainty and complexity. Ill-structuredness refers to the vagueness and unknown features of a problem, while complexity refers to the large number of problem variables and the dynamic relationships among them (Chi & Glaser, 1985; Eseryel et al., 2013; Jonassen, 1997; Kitchner, 1983; Sinnott, 1989;

Wood, 1983). According to Alexander (1997), who distinguished topical knowledge from domain knowledge, an ill-structured and complex problem situation, especially in the context of teaching, can be associated with both. The topical knowledge of a complex problem situation might be specific and deep in a particular context (e.g., identifying the complex nature of global warming and its connection to frequent wild fires in California), while its domain knowledge might cover a wide range of knowledge within a field (e.g., ecology) (Schraw et al., 1995; Shin et al., 2003). The context of the problem might be unknown or vague to a certain degree. Due to the problem's ill-defined and complex nature and the different levels of topical and domain knowledge among the students, responses to the problem could reveal diverse perspectives and conflicting solutions (Chi & Glaser, 1985; Jonassen 1997, 2000; Kitchner, 1983; Shin et al., 2003; Wood, 1983).

Many models have accounted for ill-structured problem solving as a procedural mental activity (Jonassen, 1997; Pretz et al., 2003; Sinnott, 1989). Pretz et al. (2003) identified seven primary activities: (a) recognizing the problem, (b) defining and representing the problem, (c) developing a solution strategy, (d) organizing one's knowledge about the problem, (e) allocating mental resources for solving the problem, (f) monitoring one's progress toward the goal, and (h) evaluating the solution. These problem-solving activities can typically be summarized into two main phases: (a) problem representation and (b) solution development. Concerning the first phase, Newell and Simon (1972) explained that a problem solver conceptualizes the problem space in which all of the possible conditions of a problem exist. Studies of expertise have demonstrated clear distinctions among the mentally represented problem spaces of experts and novices (Chase & Ericsson, 1982; Chi et al., 1988; Ericsson & Staszewsli, 1989; Ericsson, 2005;

Spector, 2008; Spector & Koszalka, 2004). The focus of the current study was to model the stages of building a problem space (i.e., problem representation).

2.2 Learning Progress as Changes in an Individual's Problem Space

Learning progress can be defined as a series of gradual or sudden changes in a learner's understanding (i.e., problem space) that are facilitated by instruction. The theory of mental models accounts for these cognitive changes. Mental models are cognitive artifacts that a learner constructs in order to understand a given problem situation (Anzai & Yokoyama, 1984; Craik, 1943; Greeno, 1989; Johnson-Laird, 2005a, 2005b; Kieras & Bovair, 1984; Mayer, 1989; Norman, 1983; Seel, 2003, 2004). A problem situation is mentally represented in a learner's mind when he or she is involved in the problem-solving process. Mental models depend primarily on a learner's prior knowledge (i.e., prior mental models), and they change over time (Seel, 1999, 2001). Mental model changes in a learning situation are not simple shifts without direction but transformations that lead closer to a pre-defined goal. In this way, such changes can be considered progress. Thus, progress in mental modeling involves learning-dependent and developmental transitions between preconceptions and causal explanations (Anzai & Yokoyama, 1984; Carley & Palmquist, 1992; Collins & Gentner, 1987; Johnson-Laird, 1983; Mayer, 1989; Seel, 2001, 2003, 2004; Seel & Dinter, 1995; Shute & Zapata-Rivera, 2008; Smith et al., 1993; Snow, 1990).

Mental models improve over time as a learner develops mastery in a given problem situation (Johnson-Laird, 2005a, 2005b; Seel, 2003, 2004). Progress in mental models might involve both quantitative and qualitative change. For instance, a mental model might be enlarged when a learner imports a new concept into an existing model (e.g., A sharp increase in the world population could be a major cause of global warming.). Or a mental model might fundamentally change to adapt to a new situation (e.g., The current wild fire rate is not that high because this frequency has been observed in the past. Therefore, global warming might not be associated with wild fires.).

2.3 3S Knowledge Structure Dimensions: Surface, Structural, and Semantic Dimension

Studies have shown that mental models progress as structural knowledge heightens (Johnson-Laird, 2005a, 2005b; Jonassen et al., 1993). When problem solving is defined as a mental activity that relies on structurally represented mental models (Dochy et al., 2003; Segers, 1997), the assessment of problem-solving knowledge and skills must be sensitive to the structural characteristics of the knowledge base (Gijbel et al., 2005).

Some have argued that knowledge structure consists of (a) surface, (b) structural, and (c) semantic dimensions (Kim, 2012b; Ifenthaler, 2006; Pirnay-Dummer, 2006; Spector & Koszalka, 2004). For example, Kim (2012b) demonstrated a theoretical basis for these dimensions, which had been confirmed in various studies on analogy (Gentner & Medina, 1998; Holyoak & Koh, 1987; Simon & Hayes, 1976) and linguistic comprehension (Bransford & Franks, 1971; Bransford et al., 1972; Bransford & Johnson, 1972; Fodor et al., 1974; Gentner & Medina, 1998; Katz & Postal, 1964; Kintsch & van Dijk). First, surface dimension includes descriptive information of knowledge components (i.e., concepts and their relations). This type of information is compatible with the surface level of mental models (i.e., the relevant objects and aspects of the context) (Holyoak & Koh, 1987; Simon & Hayes, 1976). According to linguistic comprehension studies, surface features, such as the relationships among nouns in a text, characterize the shape of linguistic representations when re-represented as mental models (Fodor, Bever, & Garrett, 1974; Katz & Postal, 1964).

Second, structural dimension includes the complexity and cohesiveness of a mental model as a whole. This dimension refers to a deep level of well-organized knowledge within a particular context, including underlying causal principles, key variables, and their connections (Bransford & Franks, 1971; Bransford & Johnson, 1972; Gentner & Medina, 1998; Katz & Postal, 1964; Kintsch & van Dijk, 1978).

Semantic dimension includes understanding of concepts and their relations in a knowledge structure. While surface (e.g., the number of concepts) and structural features (e.g., the complexity of a knowledge structure) refer to the generic information of an entire structure, semantic features are the individual concepts and propositional relations of a particular pair of concepts. The semantic dimension includes principle variables that emerge from information integrated into the whole structure (Kim, 2013; Bransford & Franks, 1971; Bransford et al., 1972; Bransford & Johnson, 1972; Katz & Postal, 1964; Kintsch & van Dijk, 1978).

Some studies have traced conceptual changes in response to a problem situation using parameters to quantify knowledge structure dimensions (Ifenthaler et al., 2009; Ifenthaler & Seel, 2005; Schlomske & Pirnay-Dummer, 2008). Two major limitations have been found in the literature. First, no comprehensive account of learning status has incorporated multiple knowledge structure dimensions. Second, and more seriously, many studies have analyzed learning progress based on the assumption of linear growth, resulting in an oversimplified theoretical model of learning progress.

The current study proposed that the three dimensions are not consistent across all learners but reflect how individuals develop their own mental models for solving a problem. The configurations of knowledge dimensions describing the various stages of learning progress are likely to be as diverse as the learners themselves.

3. Models of Learning Progress

Conceptual models of learning progress can be derived from studies of expertise development. Recent studies have focused on evolving expertise through learning and instruction (e.g., Kim, 2012a, 2012b; Alexander, 2003, 2004; Chi, 2006; Ericsson, 2003, 2005; Ericsson & Simon, 1980, 1993). Building on expertise development models, the current study suggests two potential candidate models that account for learning progress in complex problem solving: a three-stage model and a two-stage model.

Table 1

Models of Learning Progress in the Development of Expertise

Level	Three-Stage Model	Two-Stage Model
1	Acclimation	Misconception
2	Competence	
3	Proficiency	Conception

3.1 Three-Stage Model

Alexander (2003, 2004) introduced a multi-stage model of expertise development that focuses on academic domain learning. The Model of Domain Learning (Alexander, 2003, 2004; Alexander et al., 2004) features three stages: Acclimation, Competence, and Proficiency-Expertise. This three-stage model of expertise was empirically identified using analytic methods such as cluster analysis (Alexander & Murphy, 1998; Alexander et al., 2004). While Alexander's Model of Domain Learning theorizes that expertise develops in the long term, the current study deploys the three-stage model to explain short-term cognitive changes while solving a problem in a classroom setting. The premise that qualitatively distinct changes take place remains the same in both the short term and the long term (Siegler et al., 2009; Vosniadou et al., 2008; Vygotsky, 1934, 1978; Werner, 1957).

Table 2 describes the knowledge structure that characterizes each stage according to the current study. The "acclimation" stage refers to the initial point at which a learner becomes familiar with an unfamiliar problem. Accordingly, learners at this stage typically have a pre-structural lack of knowledge (Biggs & Collis, 1982). Another type of acclimating learner is one who has only limited and fragmented knowledge that is neither cohesive nor interconnected.

Table 2

Three-Stage Model of Learning Progress

Stage	Three Dimensions of Knowledge Structure (3S)
Acclimation	(a) All dimensions (surface, structural, and semantic) are quite dissimilar to expert models; or (b) knowledge structures have similar surface dimension to expert models but are missing some structural and semantic dimensions.
Competence	(a) Structural dimension might appear to be mastered because mental models, which consist of a small amount of contextual and abstract knowledge, are likely to look cohesive and connected; or (b) student and expert models are highly similar in semantic dimension but dissimilar in surface and structural dimensions.
Proficiency-Expertise	(a) Structural dimension shows sufficient complexity while surface dimension is adequate, but not enough to guarantee a semantic fit; (b) knowledge structures are well-featured in all dimensions (surface, structure, and semantic); or (c) a significant number of principles (semantic) create a cohesive structure (structural) but with a small number of concepts (surface).

Learners become more "competent" in understanding a problem through teaching and learning. For some competent learners, knowledge structures contain a small number of concepts and relations that appear to be structured well but still require a better sense of what is important in a particular situation. Other competent learners internalize a larger number of principles that have been taught and exhibit a good semantic dimension, however poorly organized in contextual concepts.

The last stage is "proficiency-expertise." In this stage, proficient learners, with increasing experience, construct sufficient contextual information and organize their knowledge base for a problem situation that includes some domain-specific principles. They conceptualize a sufficient problem space in all three dimensions that accommodates the real features of a given problem situation. The probability of resolving a problem markedly increases. Proficient learners sometimes represent a relatively small but efficient knowledge structure in which sufficient key concepts are well organized (good structural and semantic dimensions with a surface dimension that is lacking). These knowledge structures are in accord with the idea that experts sometimes create mental models that contain an "optimal" rather than "maximum" number of concepts and relations (Glaser et al., 1983).

3.2 Two-Stage Model

The two-stage model is another possible candidate to explain learning progress in the classroom setting where a particular problem is presented. Some studies of conceptual change provide accounts that suggest why only two stages are likely to occur.

Some have argued that conceptual change often requires shifts in an entire knowledge base (Chi, 2008; Vosniadou et al., 2008). From this point of view, a student with theoretical framework *A* (an inaccurate structure) shifts to theoretical framework *B* (an expected structure). This shift from one model to another might abruptly happen after an initially slower process. Studies have claimed that radical conceptual changes usually happen at the end of a slow process (Hatano & Inagaki, 1994; Vosniadou, 2003). Vosniadou et al. (2008) contended that the slow and gradual enrichment of knowledge is largely unconscious and that the enrichment mechanism leads to conceptual changes in the long run. For example, a recent experiment conducted by Siegler et al. (2009) confirmed that stage transitions can take place suddenly, after a given mental model has been in place for some time. They called this transition a logarithmic-to-linear shift. In their experiment, children estimated the position of a number in a line to represent numeric magnitude. Their estimations showed probabilistic patterns, which moved from being stable to approximating to the actual value. This shift occurred abruptly after applying an initial approach numerous times. In another instance, some third-grade students in a science class strongly believed that "the earth is flat" based on their naive theory. Even after instruction, they refused to accept that "the earth is spherical" and continued to justify their initial belief. Even some adults, interestingly, retain this belief (one global society still argues that "the earth is flat"; http://theflatearthsociety.org/forum/). After further instruction and additional lessons, they might begin to doubt their naive theory and, at some point, build a new one (knowledge base) that integrates new evidence and principles to support a "spherical earth."

This transition from a naive theory to an informed theory can abruptly take place at some point during a slow and gradual process of learning and instruction. From this point of view, the immature knowledge status might have two causes: (a) a pre-structural lack of knowledge (Biggs & Collis, 1982) and (b) a strong misconception. Middle stages are likely to exist, but they might have short duration and occur just before an appropriate knowledge structure takes shape. If this situation is true, measuring "middle" stages would prove difficult.

4. Methods

4.1 Participants

Participants in this investigation included 136 students and 7 professors in the area of Educational Technology. All students were enrolled in an introductory educational technology course at a large university in the southeastern United States. The course, designed to equip pre-service teachers to apply learning technologies in classroom teaching, consisted of two components: (a) introducing technology tools related to National Educational Technology Standards (NETS) and (b) designing technology-supported/enhanced lesson activities to meet NETS. The course was delivered through four sessions: one by a full-time instructor (T2) and three by three respective doctorate-level teaching assistants majoring in Instructional Technology (T1, T3, and T4). The full-time instructor had six years' teaching experience in middle schools and had taught the course for over six years. For quality control, the instructor supervised the sessions taught by the three doctoral students.

Table 3

Participant Demographics

		Total	T1	T2	T3	T4	N/A
Total		136	13	76	17	25	5
Gender	Female	113(83.1%)	11	62	14	22	4
	Male	23(16.9%)	2	14	3	3	1
Class Year	Freshman	9(6.6%)	0	6	1	2	0
	Sophomore	38(27.9%)	5	16	8	8	1
	Junior	52(38.2%)	4	31	3	13	1
	Senior	37(27.2%)	4	23	5	2	1
Course Taken	None	56(42.2%)	3	37	9	6	1
	ET-related*	8(5.9%)	2	4	0	2	0
	Other**	72(52.9%)	8	35	8	17	4
Teacher	Less than 2yrs	6(4.4%)	0	5	0	1	0
Experience	No	127(93.4%)	13	70	16	24	4
	Missing	3(2.2%)	-	-	-	-	-
Teaching	Very interested	33(24.3%)	4	20	3	5	1
Career	Interested	19(14.0%)	4	7	0	8	0
	Somewhat	32(23.5%)	2	18	6	5	1
	Not so	33(24.3%)	2	20	4	5	2
	Not at all	19(14.0%)	1	11	4	2	1

Note. *Courses that are related to Educational Technology. **Education-related courses not directly about educational technology.

Most participating students were undergraduate pre-service teachers who had considered teaching in K-12. As Table 3 shows, female students comprised 83% of the study participants, while 17% were male. The majority of students were juniors and seniors (65%), followed by sophomores (28%) and freshmen (7%). For 42% of the students, the course was their first related to educational technology, while only 6% had taken previous courses on the use of educational technology. However, over half of students had taken previous education-related courses (52%). In contrast to their coursework experience, over 93% of the students had no teaching experience. Only 38% indicated a positive interest in being a teacher even though they were all taking a course designed for pre-service teachers. The same proportion of students (38%) indicated little interest in teaching in K-12 schools.

Participants also included seven professors teaching at six different universities in the United States. They participated in the study to build a reference model for a given problem situation. In order to recruit these professors, we first created a pool of potential candidates who were members of a scholarly association related to Educational Technology and then made a short list of professors based on pre-set criteria: (a) professors in Instructional Technology or related fields; (b) professors teaching a course titled “Instructional Design” or “Technology Integration in Learning”; (c) professors who research technology-integration in classroom learning; and (d) professors whose doctorates were received at least three years ago. Seventeen professors were invited to participate in the study, and seven professors were willing to contribute their expertise to the research.

4.2 Data Collection

4.2.1 Problem Task

This study collected participants’ responses to an ill-structured and complex problem task, a simulated situation that participants were asked to evaluate. The simulated situation was a failed attempt to adapt a technology (i.e., tablet PC) to classroom teaching. Designed to elicit participants’ knowledge, the questions asked them to describe, in explicit detail, the concepts, issues, factors, and variables likely to have contributed to the result of the project: introducing tablet PCs had very little effect on the instructional practices employed in the classroom (see Appendix A).

In the third week of the semester, the instructors asked participating students to respond to the problem situation as an in-class problem-solving activity. The participants wrote their responses using natural language, which was required for two reasons. First, using natural language enables individuals to verbalize their understanding of a problem situation and some feasible solutions (Kim, 2013). Second, using natural language is more likely to provide a reliable foundation whereby a descriptive knowledge representation can be elicited (Kim, 2012b, 2013; Pirnay-Dummer et al., 2010).

4.2.2 Reference Model

Using Delphi survey procedures, seven professors established a reference model against which student models could be compared (Goodman, 1987; Hsu & Sandford, 2007; Okoli & Pawlowski, 2004). The Delphi survey included three rounds to refine the reference model (see Table 4). In the first round, the professors created their own responses to the problem. These initial responses were consolidated for use in the second round and also used as individual cases for analysis. Next, the consolidated responses and a list of identified concepts were sent to the panel. The experts were asked to comment on the listed statements and concepts and rank them. After gathering the second-round surveys, the researcher created a final list of ranked statements and concepts. Based on the summary, a draft of the reference model was created. In the final round, the results of the second survey were sent to the panel and revised according to their comments, where necessary. The outcome of these procedures was a written reference model containing 23 key concepts identified by the panel (see Appendix B). The reference model was believed to include most of the concepts and relations required to form a knowledge structure representing the problem situation.

Table 4

Delphi Procedure

Round	Activity
R1. Brainstorming	• Collect and consolidate all responses from experts
R2. Narrowing Down/ Ranking	• Send refined final version of consolidated lists, including statements and used concepts • Ask experts to add comments if they disagree with or have different opinions about a statement • Ask experts to rank key statements and concepts
R3. Refinement	• Send each panelist ranked statements and concepts summarized by the investigators • Ask for revision of judgments or specification of reasons for remaining outside the consensus

Note. Adapted from Kim (2013, p. 960).

4.3 Data Analysis

4.3.1 Procedure

This study attempted to validate the models of learning progress in problem solving using empirical data gathered from written responses. This process entailed transforming lexical data into numerical data that described the knowledge structure dimensions and then analyzing the transformed data using a particular set of analytic techniques.

Data transformation was conducted using a three-step concept mapping technique (Kim, 2012a; Kim, 2013; Clariana, 2010; Curtis & Davis, 2003; Narayanan, 2005; Novak & Canãs, 2006; Spector & Koszalka, 2004; Taricani & Clariana, 2006): (a) elicit concepts and relations in a text, (b) construct concept maps drawn by graph theory (Rupp et al., 2010a; Schvaneveldt et al., 1989; Wasserman & Faust, 1994), and (c) compare the concept maps to a reference model.

The first step was to elicit judgments about concept relations in students' written responses to the problem. This study made judgments in compliance with Kim's (2012a, 2013) Semantic Relation (SR) approach, which obtains semantic relations from the syntactic structure of a written response and then creates a concept map that better represents a student's internal mental representation. The second step was to construct concept maps. This study used the network analysis software NetMiner (http://www.netminer.com) to draw concept maps and generate concept map parameters. This tool automatically rendered the concept map parameters, based on indicators suggested for educational diagnostics. The final step was to compare student concept maps to a reference model. This comparison yielded similarity measures that were later used for model validation. Similarity measures were calculated using a similarity analysis tool developed by Kim (2012a) using the C++ programming language. For a more detailed account of this concept map method, see Kim (2013).

The similarity measures were continuous variables ranging from 0 to 1 (see details in the following section). The similarity measures were first reviewed using SPSS 21 (http://www-01.ibm.com/software/analytics/spss/) to determine whether they violated the multivariate normality assumption and whether there were outliers. Next, the validated data was analyzed using Exploratory Factor Analysis (EFA) in order to discover any relations between the similarity measures and the 3S knowledge structure dimensions (i.e., latent factors). Lastly, the models were validated with the Continuous Log-Linear Cognitive Diagnostic Model (C-LCDM) method (Bozard, 2010). R software (http://www.r-project.org/) was used to run EFA, and M-Plus version 6.0 (https://www.statmodel.com/) was used for C-LCDM analysis.

4.3.2 Similarity Measures as Indicators of Levels of Understanding

The human mind is not easily observed, but cognitive activity can be indirectly inferred from externalized representations, such as written responses to a problem situation. Drawing on the aforementioned concept mapping technique, participants' written responses were transformed into concept maps (see examples in Figure 2), whereby a list of parameters might be obtained. In terms of graph theory, Ifenthaler (2010) introduced parameters that might be used to diagnose knowledge representation in an educational setting. After reviewing these parameters, using network analysis methods (Coronges et al., 2007; Ifenthaler, 2010; Wasserman & Faust, 1994), the current study proposed a new set of parameters that could demonstrate knowledge structure dimensions (see Table 6). For example, the parameter "average degree" indicates the average number of relations (links) to or from a concept (node), ranging from 0 to g-1 (g being the total number of nodes). As average degree increases, the knowledge structure is considered more complex.

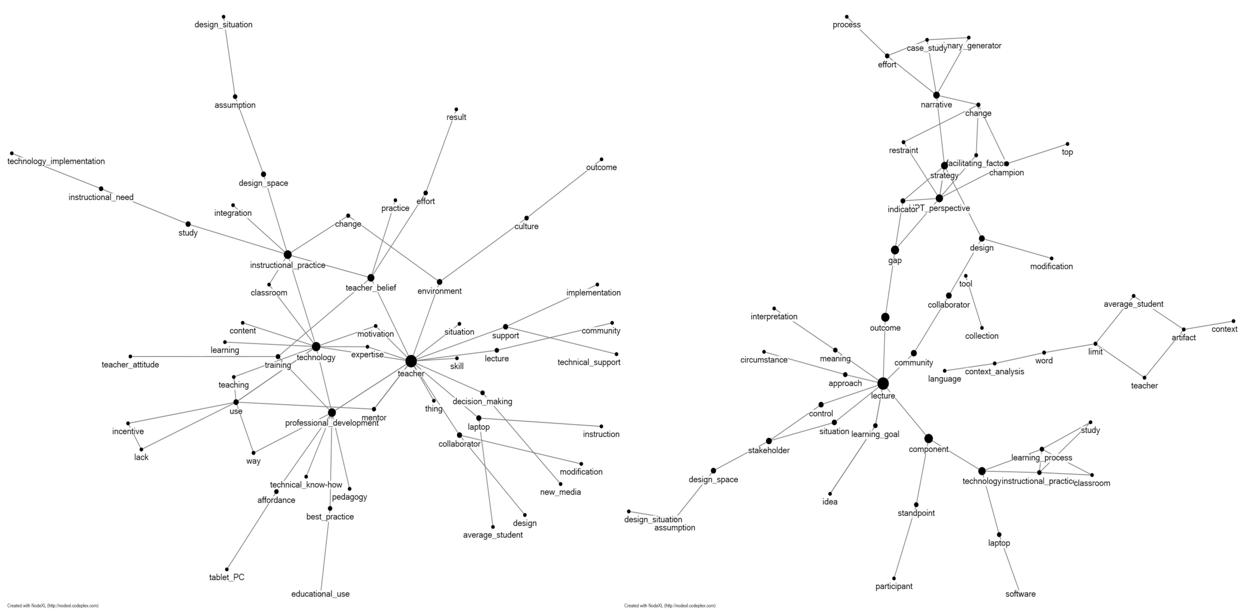

Figure 2. Examples of concept maps elicited from participants' written responses.

Table 5

Similarity Measures

Similarity Measure	Definition	Parameters Compared**	
		Technical Definition	Operationalization
1. Number of concepts	Compare the number of concepts (nodes) in two models	The total number of nodes (vertices)	The overall number of concepts[a]
2. Number of relations*	Compare the number of links (edges) in two models	The total number of links (edges)	The overall number of relations of paired concepts[a]
3. Average Degree	Compare the average number of degrees in two models	The average number of links of a node, ranging from 0 and *g*-1 (*g* being the total number of nodes)	As the number of incoming and outgoing relations grows, the complexity of the cognitive structure is considered higher.[a]
4. Density of graphs	Compare the density of two models	The density of a graph is the proportion of possible lines that are actually present in the graph.	The density of a concept map indicates how cohesive it is.
5. Mean Distance	Compare the mean distances in two models	The average geodesic*** distance between any pair of nodes in a network	Indicates how close the concepts are to one another.
6. Diameter*	Compare the largest geodesics in two models	The length of the largest geodesic between any pair of nodes (1 to *g*-1)	Represents how broad the understanding of a domain is[a]
7. Connectedness	Compare the ratios of paired nodes that reach each other in each graph	Ratio of pairs, nodes that can reach each other in the graph.	Describes the extent to which the concepts are connected[a]
8. Subgroups	Compare the number of cohesive subgroups in each graph	Subsets of actors among which there are relatively strong, direct, intense, frequent, or positive ties.	The more complex a cognitive structure, the more subgroups will intermediate the entire network[a]
9. Concept Matching*	Compare semantically identical concepts, including contextual and principle variables	Qualitative comparison	
10. Propositional Matching*	Compare fully identical propositions (edges) between two concept maps	Qualitative comparison	
11. Balanced Semantic Matching*	Compare the balances calculated by dividing Propositional Matching by Concept Matching	Based on the values of two similarities: Concept and Propositional Matching	

Note. *These similarity measures are adopted from Pirnay-Dummer & Ifenthaler (2010). **Similarity measures are calculated by comparing the parameters that are suggested in Wasserman and Faust (1994). ***The shortest path between two nodes is referred to as a geodesic. [a]These parameters are also introduced by Ifenthaler (2010).

Evaluation of a student's concept map is often accomplished by comparing it to a reference model, which is usually elicited from an expert's response (Coronges et al., 2007; Curtis & Davis, 2003; Goldsmith & Kraiger, 1997; Taricani & Clariana, 2006). This comparison assesses similarity measures by overlaying network patterns with the concept map parameters (Coronges et al., 2007; Monge & Contractor, 2003). Similarity measures at each measurement occasion can indicate how close a learner model is to a reference model. The current study proposed eleven similarity measures applicable to the study of cognitive change (see Table 5).

To compare parameters, two types of similarity formulas were applied: (a) numerical similarity and (b) conceptual similarity. For most of the similarity measures (i.e., number of concepts, number of relations, average degree, density of graphs, mean distance, diameter, connectedness, and subgroups), the numerical similarity formula was used because the parameters were numerical values.

On the whole, a similarity formula assumes that each half of a pair is equally significant. In the case of a concept model comparison, the reference model and student model are not equivalent in terms of maturity. A reference model acts as a standard toward which a student model is expected to progress. A reference model is likely to contain a greater number of concepts and relations and is comprised of a larger and more complex knowledge structure than a novice model (Chi, Glaser, & Farr, 1988; Spector & Koszalka, 2004). Thus, a modified formula was used for those measures because an optimal value, rather than a greater value, indicated a good condition (see Table 5). When f_1 was less than f_2 ($f_1 < f_2$), the original numerical similarity formula was used so that

$$s = 1 - \frac{|f_1 - f_2|}{\max(f_1, f_2)},$$

where f_1 is the frequency of a student model and f_2 is the frequency of a reference model. The similarity ranged from 0 to 1, $0 \leq s \leq 1$. Otherwise, if f_1 was not less than f_2 ($f_1 \geq f_2$), the similarity value was set to "1" because the student value was greater than the reference value, indicating that the student model exceeded the reference model in the relevant criteria.

For the three similarity measures (concept, proposition, and balanced semantic matching score), conceptual similarity was applied because these measures concern the proportion of fully identical elements between two concept maps. The conceptual similarity method used in this study applied Tversky's (1977) similarity formula, which assumes that the similarity of object A and object B is not merely a function of the features common to A and B but relies on the unique features of each object. As in the case of numerical similarity, an adjustment was made. Just as a picture resembles an object rather than the inverse, a student model resembles the reference model, which is more salient. In this asymmetric relation, the features of the student model are weighted more heavily than the reference model features (Colman & Shafir, 2008; Tversky & Shafir, 2004). When the conceptual similarities were calculated using Tversky's (1977) formula, in this study α was weighted more heavily than β ($\alpha = 0.7$ and $\beta = 0.3$).

$$s = \frac{f(\mathrm{A} \cap \mathrm{B})}{f(\mathrm{A} \cap \mathrm{B}) + \alpha \cdot f(\mathrm{A} - \mathrm{B}) + \beta \cdot f(\mathrm{B} - \mathrm{A})}$$

Each of the similarity measures was assumed to be associated with at least one of the three dimensions (i.e., 3S: Surface, Structural, and Semantic) of knowledge structure. Presumably, the similarity measures demonstrate how close a student's knowledge structure features are to the reference model.

4.3.3 Validation Method: Continuous Log-Linear Cognitive Diagnostic Model (C-LCDM)

This study adapted Latent Class Models (LCM) to identify the stages of learning progress. The main assumption was that the qualitative stages of learning progress can be labeled "latent classes" due to their psychometric characteristics. LCMs consist of corresponding latent classes associated with the other stages so that on any given occasion of measurement, each individual has an array of latent class memberships (Collins & Cliff, 1990; Collins et al., 1994; Heinen, 1996; Kaplan, 2008; Rost & Langeheine, 1997).

As a special case of constrained LCM, Log-linear Cognitive Diagnostic Model (LCDM) allows latent class models to place linear restrictions on the log-linear parameters when both observed variables and latent predictor variables are categorical, especially dichotomous (Agresti, 2007; Rupp et al., 2010b). According to Templin (2004), latent predictor variables, often called *attributes*, "are the atomic components of ability, the specific skills that together comprise the latent space of general ability" (p. 8). Attribute variables have a binary continuum: "mastery" vs. "non-mastery." An individual attribute is considered "mastery" when its probability is higher than 0.5 (Rupp et al., 2010b). LCDMs are used to determine whether an individual masters the attributes required to answer items correctly. The attributes that have or have not been mastered constitute an individual's mastery profile, which ultimately defines his/her latent class membership (Henson et al., 2009).

The current study posited that the 3S dimensions of a student's knowledge structure are knowledge attributes. The status of each knowledge attribute is either "mastery" or "non-mastery." The 3S attributes of knowledge structure meet the assumption of LCDM. However, this study did not directly use LCDM because it used continuous observed variables (i.e., similarity measures) instead of dichotomous variables.

Bozard (2010) proposed the Continuous LCDM in order to accommodate continuous models. She defined C-LCDM using the following equation, which is similar to the CFA model:

$$x_{ij} = \lambda_{i0} + \lambda_{i1}\alpha_{j1} + \lambda_{i2}\alpha_{j2} + \lambda_{i12}\alpha_{j1}\alpha_{j2} + e_{ij}, \qquad (1)$$

where x_{ij} symbolizes the response to the observable variable i by person j; λ_i indicates factor loading (i.e., attribute loading) to the item, which takes three types of the following values: intercept = 0, main effect = 1, two-way interactions = 2, etc.; α_{jk} symbolizes latent variable value (i.e., attribute value) for individual j and latent variable k (k = 2 in equation 1 above); and e_{ij} represents the uniqueness of the individual in observed indicator variables (residuals).

C-LCDMs contain a set of attribute mastery profiles in that the set of k latent attributes can be considered a latent class model with 2^k classes (Henson et al., 2009; Templin, 2004). In the context of modeling learning progress, the three dimensions of knowledge structure (i.e., k = 3 latent attributes: surface, structural, and semantic) are attributes in C-LCDMs. Each dimension (attribute) of knowledge structure can be profiled as mastered/obtained (1) or non-mastered/not obtained (0). Those mastery profiles provide stages (i.e., 2^3 = 8 latent classes) of learning progress at which an individual might be classified. Based on the three Models of Learning Progress discussed earlier, these eight latent classes are associated with the stages of each model (see Table 6).

Table 6

Class-to-Profile Table

Latent Class	3S Attribute			Stages of Learning Progress	
	S1	S2	S3	Three-Stage Model	Two-Stage Model
Class 1 (C_1)	0	0	0	Acclimation	Misconception
Class 2 (C_2)	0	0	1	Competence	-
Class 3 (C_3)	0	1	0	Competence	-
Class 4 (C_4)	0	1	1	Proficiency	Conception
Class 5 (C_5)	1	0	0	Acclimation	Misconception
Class 6 (C_6)	1	0	1	-	-
Class 7 (C_7)	1	1	0	Proficiency	Conception
Class 8 (C_8)	1	1	1	Proficiency	Conception

Note. 3S attributes involve the three dimensions of knowledge structure: Surface (S1), Structural (S2), and Semantic (S3). 0 = absent/non-mastered and 1 = present/mastered.

For instance, class 4 has a mastery profile (i.e., S1: 0, S2: 1, and S3: 1) that is matched to the proficient learner in the three-stage model, and conception in the two-stage model. These matches were determined by theoretical assumptions about knowledge status at each level. For example, a proficient learner represents a relatively small number of concepts and relations (S1: 0 = surface feature is absent) but an efficient knowledge structure (S2: 1 = structure feature is present) in which sufficient key concepts are well-structured (S3: 1 = semantic feature is present).

Table 7

Descriptive Statistics of Similarity Measures

	N	Minimum	Maximum	Mean	SD
M1. Concept	140	0.04	0.96	0.35	0.18
M2. Relation	140	0.02	0.98	0.31	0.20
M3. Average Degree	140	0.42	1.00	0.82	0.14
M4. Density	140	0.05	0.96	0.40	0.19
M5. Mean Distance	140	0.28	0.99	0.73	0.17
M6. Diameter	140	0.14	1.00	0.75	0.22
M7. Connectedness	140	0.15	1.00	0.73	0.25
M8. Subgroups	140	0.00	1.00	0.30	0.20
M9. Concept Matching	140	0.05	0.57	0.24	0.08
M10. Propositional Matching	140	0.00	0.28	0.05	0.04
M11. Balanced Matching	140	0.00	0.68	0.18	0.15

5. Results

5.1 Descriptive Statistics

Outlier analysis for the similarity measures was conducted first. As a result, three student cases were eliminated, resulting in a sample size of 140, including the seven experts. Then, we examined the descriptive statistics and correlations of the eleven similarity measures as continuous variables ($0 \leq s \leq 1$). As Table 7 presents, most similarity measures showed biased distributions. The similarity measures M9 to M11 were distributed in the lower area of the similarity band, ranging from 0 to 0.68, with lower means ranging from 0.05 to 0.24, while the distributions of M3, M5, M6, and M7 ranged from 0.15 to 1.00, with means ranging from 0.73 to 0.82. M1, M2, M4, and M8 had means between 0.31 and 0.40. These descriptive results suggest that the similarity measures likely indicate different constructs, which the current study assumes to be the different dimensions of knowledge structure.

Table 8

Correlations of the Similarity Measures

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
M1	1										
M2	.97*	1									
M3	.60*	.71*	1								
M4	.95*	.85*	.39*	1							
M5	.64*	.61*	.65*	.59*	1						
M6	.65*	.62*	.62*	.63*	.94*	1					
M7	.01	.16	.51*	-.20*	.36*	.21*	1				
M8	.92*	.92*	.63*	.85*	.71*	.72*	.14	1			
M9	.54*	.56*	.49*	.46*	.43*	.46*	.06	.50*	1		
M10	.38*	.43*	.37*	.27*	.21*	.22*	.14	.37*	.61*	1	
M11	0.16	.21*	.24*	0.08	0.08	0.08	0.13	.15	0.26*	.86*	1

n = 140. *Correlation is significant at the 0.05 level (2-tailed).

Next, as shown in Table 8, correlations among the similarity measures were calculated. Very high correlations ($r > .9$) were identified among variables M1, M2, M4, and M8. Strong correlations also emerged between two paired variables, M4 and M8 ($r = 0.85$) and M10 and M11 ($r = 0.86$). These high correlations might provoke multicollinearity concerns for a general linear model, such as multivariate regression; however, they were less problematic for C-LCDM because LCDMs allow multiple items (i.e., observed variables) to measure the same set or similar sets of attributes (Rupp et al., 2010b). Consequently, we retained all 11 measures for further analysis. The overall Kaiser-Meyer-Olkin (KMO) value was 0.7, within the range of good (between 0.7 and 0.8), even though two individual measures were below the recommended value of > 0.5 (M7 = 0.42 and M14 = 0.42). Barlett's test of sphericity, x^2 (55) = 2570.92, $p < 0.01$, indicated that correlations between variables were significantly large; therefore, factor analysis was appropriate.

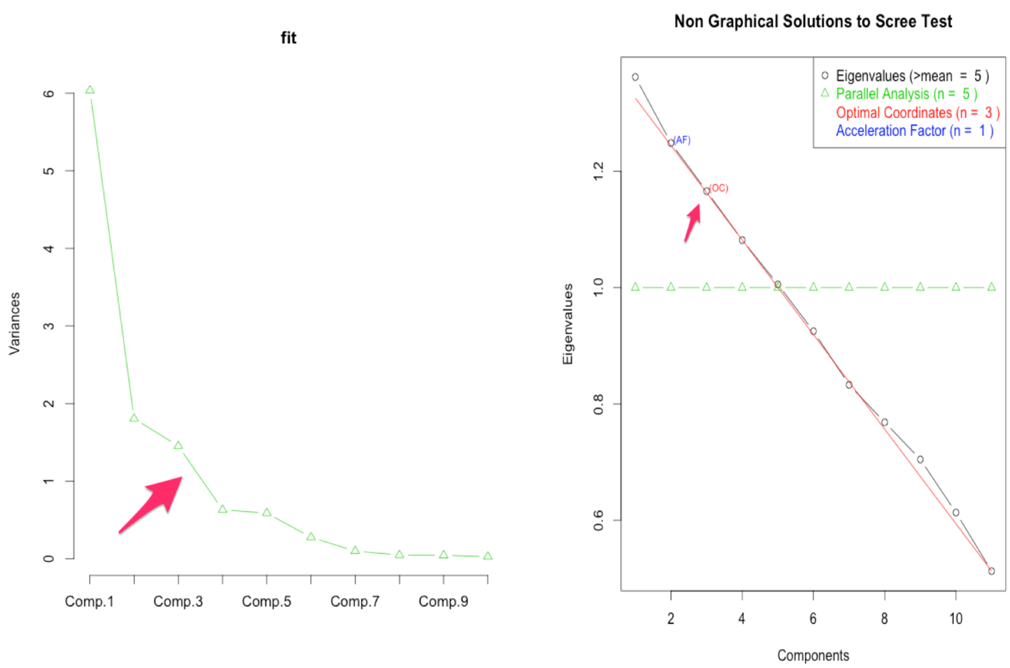

Figure 3. Number of factors to extract by Scree plots.

5.2 The Relations between Similarity Measures and 3S Knowledge Structure Dimensions

We anticipated that the eleven similarity measures would be associated with three latent factors: the surface, structural, and semantic dimensions of knowledge structure. In order to test this assumption, we ran an initial analysis to determine the number of factors we had to extract from the data. The principle component analysis calculated eigenvalues for each component and produced two types of scree plots (see Figure 3).

Three factors on the left graph had eigenvalues over Kaiser's criterion of 1, and both graphs suggested that these three factors were above the cut-off point of inflexion, explaining 76% of the variance. Consequently, three factors were retained for final analysis.

Table 9 shows the factor loadings after oblique rotation. The similarity measures that clustered on the same factors suggest that factor 1 represents "surface," factor 2 "structural," and factor 3 "semantic." For instance, "M4. Density," clustered on factor 1 (surface), is defined in graph theory as the proportion of existing links to all possible lines in the graph (Wasserman & Faust, 1994). Density is associated with the number of existing links (i.e., a surface feature). Amongst the similarity measures, three measures ("M3. Average Degree," "M7. Connectedness," and "M9. Concept Matching") loaded on multiple factors. For example, the variance of "M3. Average Degree" was explained by surface and structural features. Interestingly, "M7. Connectedness" was negatively related to "surface" while positively associated with "structural." In addition, communality analysis verified sample-size ($N = 140$) adequacy with values over 0.5, except M7, based on the suggestion that sample sizes between 100 and 200 are large enough to conduct factor analysis as long as there are relatively few factors and the communality values are above 0.5 (Guadagnoli & Velicer, 1988).

Table 9 presents factor loadings, the absolute values of which were all over 0.3, a reliable cut-off point. However, Stevens (2002) claimed that the significance of factor loading is determined by the sample size and the variance explained by a particular factor. According to his suggestion, with a sample size of 140, factor loadings should be greater than 0.4 and explain around 16% of the variance in a similarity measure. When we applied the cut-off value of 0.4, all but three factors had single loadings on one item.

Table 9

Summary of Exploratory Factor Analysis

		Oblique Rotated Factor Loadings		
Similarity Measure	Communality	Surface	Structure	Semantic
M1. Concept	99.50	1.01		
M2. Relation	94.71	0.96		
M3. Average Degree	51.72	(0.31)	0.42	
M4. Density	90.47	1.00		
M5. Mean Distance	99.50		0.96	
M6. Diameter	89.90		0.86	
M7. Connectedness	22.85	(-0.35)	0.55	
M8. Subgroups	88.17	0.82		
M9. Concept Matching	49.65	(0.30)		0.46
M10. Propositional Matching	99.50			1.01
M11. Balanced Matching	77.99			0.95

Note. Factor loadings in parentheses are excluded when the 0.4 cut-off value is applied.

Having found that three factors (surface, structural, and semantic dimension) were associated with the similarity measures, we created confirmatory models known as Q-matrix models. Like all LCDMs, C-LCDM requires a substantive theoretical model whereby researchers can interpret statistical classifications as meaningful latent classes. The Q-matrix is a hypothetical model that defines a limited relationship between a set of attributes and a set of test items (Templin, 2004). In the case of the current study, the value of each similarity measure is dependent on mastering the designated attributes (required knowledge dimensions).

Table 10 describes the hypothesized relationships between the 11 similarity measures and the three dimensions of knowledge structure (i.e., attributes). In the table, an attribute is coded “1” when it is required to raise the value of the measure. Two Q-matrices were designed, one for each cut-off point. Using the 0.4 cut-off value, the mastery associations in parentheses were excluded (i.e., for measures M3, M7, and M9). For convenience, this Q-matrix model was called “Main effect only (Model 1).” Using the more relaxed 0.3 cut-off value, the Q-matrix had multi-loading on measures M3, M7, and M9 and was called “Interaction effect (Model 2).” Also noteworthy is that the patterns of mastering attributes defined an individual’s latent classes (see Table 6).

Table 10

Similarity Measures Associated with Attributes of Knowledge Structure (Q-Matrix)

Similarity Measure	3S Attribute		
	Surface (S1)	Structure (S2)	Semantic (S3)
M1. Concept	1	0	0
M2. Relation	1	0	0
M3. Average Degree	(1)	1	0
M4. Density	1	0	0
M5. Mean Distance	0	1	0
M6. Diameter	0	1	0
M7. Connectedness	(1)	1	0
M8. Subgroups	1	0	0
M9. Concept Matching	(1)	0	1
M10. Propositional Matching	0	0	1
M11. Balanced Matching	0	0	1

Note. According to the factor loadings (see Table 9), the mastery associations in parentheses (i.e., for measures M3, M7, and M9) were excluded (Cut-off value: 0.4)

5.3 Investigation of the Models of Learning Progress Using C-LCDM

Model estimation was conducted for Main effect only (Model 1) and Interaction effect (Model 2) in order to obtain model-fit indices. Then, in order to select the best model, AIC and

BIC indices were compared because the likelihood ratio test tends to be unreliable when using mixture analysis.

As Table 11 shows, Model 2, a complex model, had slightly better AIC and BIC values, that were little smaller than the Model 1 values. However, because the differences were so small, we decided to use the simpler model (i.e., Model 1).

Table 11

Comparison of Relative Fit of C-LCDMs

Model	Log-likelihood	AIC	BIC
Main Effect Only (Model 1)	1078.495	-2076.990	-1959.324
Interaction Effect (Model 2)	1093.882	-2095.763	-1960.448

Note. "Model 2" allowed interaction effects for M3 and M9. *Akaike's information criterion* (AIC); *Bayesian information criterion* (BIC).

M-plus reported the estimated posterior probabilities and the most likely latent class for each respondent. Table 12 describes the posterior probability results, including the final counts and the proportions of the latent classes.

Three latent classes were identified: 1, 3, and 7. Based on the learning progress models, each class was matched to the appropriate stage ("acclimation," "competence," and "proficiency" from the three-stage model, and "misconception" and "conception" from the two-stage model). Although the two-stage model was explained with the identified latent classes, the three-stage model generated the most matches, including the middle stage, "competence." Accordingly, we concluded that the three-stage model better accounts for the multiple stages of expertise development.

Table 12

The Estimated Final Class Counts and Proportions

Latent Class	Q-Matrix			Estimated Classification (Count)	
	S1	S2	S3	Model 1	Model 2
Class 1 (C_1)	0	0	0	0.288(38)	0.274(37)
Class 2 (C_2)	0	0	1	-	-
Class 3 (C_3)	0	1	0	0.427(62)	0.418(60)
Class 4 (C_4)	0	1	1	-	-
Class 5 (C_5)	1	0	0	-	-
Class 6 (C_6)	1	0	1	-	-
Class 7 (C_7)	1	1	0	0.285(40)	0.308(43)
Class 8 (C_8)	1	1	1	-	-

Stages of Learning Progress		
Latent Class	Three-Stage Model	Two-Stage Model
Class 1 (C_1)	**Acclimation**	**Misconception**
Class 2 (C_2)	Competence	
Class 3 (C_3)	**Competence**	
Class 4 (C_4)	Proficiency	Conception
Class 5 (C_5)	Acclimation	Misconception
Class 6 (C_6)	-	-
Class 7 (C_7)	**Proficiency**	**Conception**
Class 8 (C_8)	Proficiency	Conception

Table 13 lists the parameter estimates of the Main Effect Only (Model 1). These parameter estimates could be used for standardizing a C-LCDM if the sample size were much larger than the one in this study. To illustrate how these parameter values might be applied, we calculated the parameter estimates associated with M1, concept similarity. M1 measures only one surface dimension. The intercept for M1 was estimated to be 0.420, meaning that an individual who has not mastered the surface dimension has an average correct response of 0.420. The main effect for M1 for the surface dimension was estimated to be 0.167, meaning that an individual who has mastered the attribute has an average correct response of 0.587 (0.420+0.167).

However, because this study piloted the C-LCDM with a small number of samples (*N*=140), the interpretation of the parameter estimates should be limited to this study.

Table 13

Parameter Estimates for Main Effect Only (Model 1)

Parameter	Estimate	Standard Error
M1_intercept	0.420	0.037
M1_main effect of surface	0.167	0.020
M2_intercept	0.384	0.038
M2_main effect of surface	0.178	0.017
M3_intercept	0.772	0.015
M3_main effect of structure	0.109	0.011
M4_intercept	0.462	0.038
M4_main effect of surface	0.157	0.025
M5_intercept	0.671	0.019
M5_main effect of structure	0.145	0.013
M6_intercept	0.676	0.025
M6_main effect of structure	0.179	0.019
M7_intercept	0.678	0.026
M7_main effect of structure	0.124	0.027
M8_intercept	0.377	0.041
M8_main effect of surface	0.173	0.027
M9_intercept	0.664	0.006
M9_main effect of semantic	0.423	0.000
M10_intercept	0.522	0.004
M10_main effect of semantic	0.476	0.000
M11_intercept	0.645	0.013
M11_main effect of semantic	0.465	0.000

Note. M1 (Concept), M2 (Relation), M3 (Average Degree), M4 (Density), M5 (Mean Distance), M6 (Diameter), M7 (Connectedness), M8 (Subgroups), M9 (Concept Matching), M10 (Propositional Matching), and M11 (Balanced Matching).

Lastly, based on the three-stage model, we investigated the associations between the similarities and the three identified latent classes: class 1 (Acclimation); class 3 (Competence); and class 7 (Proficiency). As Table 14 shows, a series of Spearman rank-order correlations indicated that the ordered classes had a significant positive relationship with all similarity measures. Group mean difference tests proved that there were significant differences among the groups in all similarity measures except M11 (Balanced matching) (F (2, 130) = 1.97, p > 0.05).

Table 14

Means, Standard Deviations, Spearman's rho, and Group Difference Statistics

	Class 1 Mean (SD)	Class 3 Mean (SD)	Class 7 Mean (SD)	$r_s(2)$	F	Sig.
M1. Concept	.194 (.080)	.290 (.070)	.587 (.140)	.829**	178.69	.000
M2. Relation	.130 (.057)	.252 (.079)	.563 (.161)	.873**	180.81	.000
M3. Average Degree	.655 (.101)	.850 (.092)	.923 (.079)	.725**	91.43	.000
M4. Density	.270 (.122)	.326 (.087)	.618 (.154)	.691**	102.02	.000
M5. Mean Distance	.515 (.108)	.777 (.104)	.870 (.084)	.779**	134.19	.000
M6. Diameter	.485 (.157)	.802 (.138)	.929 (.112)	.756**	109.61	.000
M7. Connectedness	.547 (.282)	.825 (.197)	.759 (.185)	.266**	19.18	.000
M8. Subgroups	.119 (.065)	.256 (.070)	.551 (.164)	.880**	175.92	.000
M9. Concept Matching	.199 (.075)	.245 (.055)	.274 (.087)	.332**	11.20	.000
M10. Propositional Matching	.032 (.036)	.042 (.034)	.066 (.059)	.223**	6.59	.002
M11. Balanced Matching	.154 (.170)	.171 (.139)	.217 (.138)	.175*	1.97	.144

Note. *$p \leq .05$. **$p \leq .01$. $r_s(2)$ denotes Spearman's rank order correlations with 2 degree of freedom. Standard deviations appear in parentheses next to means.

6. Conclusion

6.1 Discussions and Implications

The primary goal of this study was two-fold. One was to model different levels of learning progress in the context of teaching and learning when solving ill-structured, complex problems. Theoretical discussions about expertise development underpin the models of learning progress, including two-stage and three-stage models. The other goal was to investigate a model of learning progress that better fits the data obtained from 133 students and 7 expert professors. The validation analysis using C-LCDM demonstrated that the data contained three latent classes that exactly corresponded to the three-stage model (see Figure 4): (a) acclimation, (b) competence, and (c) proficiency.

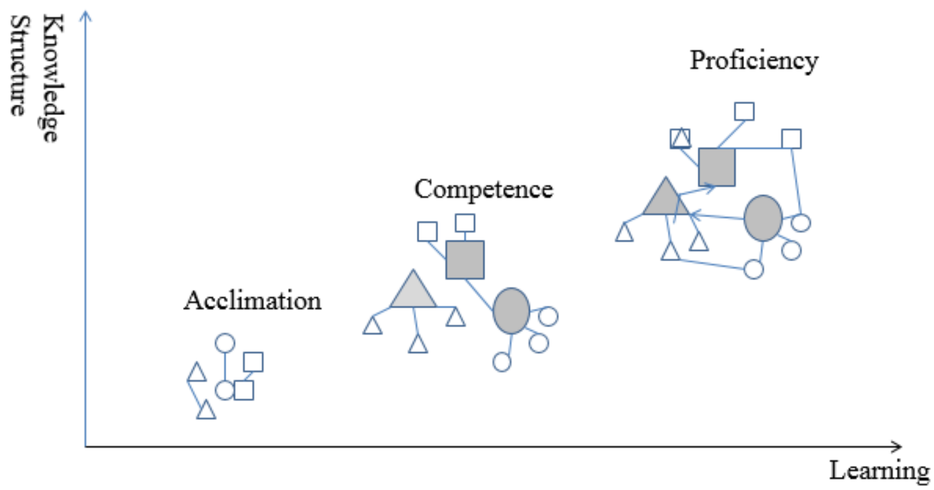

Figure 4. The measurable three stages of learning progress modified from Kim (2012b).

Although the two-stage model was also explained by these three latent classes, the study suggests that the three-stage model provides a better framework for measuring learning progress in complex problem solving. In the three-stage model, the middle stage overcomes the limitation associated with contrasting extremes, such as conception (i.e., experts) and misconception (i.e., novice).

The three levels of learning progress explain cognitive changes over a short period of practice, such as in-class problem solving, that correspond to Alexander's (2003, 2004) three levels of domain learning that develop over a relatively long period of time (i.e., featured acclimation cluster, competence cluster, and proficiency cluster). Interestingly, whether in the short term or the long term, the three stages suggest that learners are likely to experience three qualitatively distinct levels in learning and development (Siegler et al., 2009; Vygotsky, 1934/1978; Werner, 1957).

The current study observed two key conditions that relate to the stages of learning progress: (a) teaching and learning and (b) prior knowledge. The research context implies that students' prior experience played an essential role in determining their levels of understanding. Since data were collected early in the semester, participants in this study responded to the problem without instruction in related content. Moreover, all participants were undergraduates, most of whom were non-education majors. In this limited-instruction condition, association with teaching experience showed a significant chi square value, yet there was no significant association between levels of learning progress and student demographics: school year, interest in becoming a teacher, and prior coursework. Still, we withheld its generalized interpretation because some of the expected cell values were less than five frequencies (see Appendix C).

We also had to consider the ill-structured and complex nature of a given problem (e.g., its unknown elements and high number of interactions between associated factors; see Choi & Lee, 2009; Eseryel, et al., 2013; Jonassen, 2014; Kim, 2012b). The nature of the problem explains the reason there was no other latent class pertaining to the proficiency stage. For example, due to the problem's unknown elements and high complexity, the participants rarely conceptualized a knowledge structure featuring all dimensions (i.e., Class 8).

Interestingly, all seven experts were identified as proficient, but unexpectedly, their knowledge structure did not perfectly match the reference model in all three dimensions. The current study began with the assumption that experts in a domain would build similar understandings of a problem despite its ill-structuredness and high complexity. However, the results showed that their proficiency was built on their own belief, forming somewhat different knowledge structures that emphasized some concepts and facts in the reference model. Although they constructed a reference model in collaboration with each other, the individual expert models did not include most of the internal and external factors associated with a given problem situation (c.f., Jonassen, 2014).

These findings imply that we might observe different sets of latent classes within the three-stage model. For example, complex yet well-structured problems are likely produce common and convergent answers (Bogard, et al., 2013; Jonassen, 2014; Shin et al., 2003). These answers might represent two types of proficiency: (a) latent class 4 (a cohesive knowledge structure with a significant number of principles (semantic) but a small number of concepts (surface)); or (b) latent class 8 (a well-featured knowledge structure at all levels (surface, structural, and semantic)).

Lastly, the three-stage model of learning progress and proposed measurement could guide the development of technology-enhanced assessment applications. Current demands for student-centered learning require prompt and precise understanding of student changes, both cognitive and non-cognitive (Lee & Park, 2007). Enhanced knowledge about students is critical to helping students engage in their own learning. For example, one such application might be an in-class response system. If a teacher projected a problem case to the whole class, the students could submit written responses to the system. The system could then analyze the students' knowledge structures, levels of understanding, types of expertise, and missing or inaccurate concepts. Finally, all assessment results could be delivered to a teacher in the form of organized and visible data. Another example might be an advanced Intelligent Tutoring System (ITS). Using an ITS for science education, students could try to identify the complex nature of global warming and how it might be affecting recent wild fires in California and then design an environmental structure to resolve that problem. In this situation, students could explore a virtual space and conduct research guided by an agent. Once the student verbally explained his or her understanding of a complex problem situation to an agent, the response could be assessed by the embedded technology. Drawing on the assessment results, the agent could automatically deliver formative feedback and adaptive instructional support (e.g., different instructional modes and levels of explanation depending on each student's level of learning progress).

6.2 Limitations and Suggestions

The suggested methods and findings of this study open pathways to future research. Some potential avenues are explained below. First, as a basis for detecting and validating changes in conceptual development, these models and methods are applicable to a wide range of areas: conceptual change in a problem-solving situation, linguistic comprehension, evaluation of scientific argumentations, expertise modeling, and longitudinal studies of learning progress. For example, investigating longitudinal changes in learning can help us evaluate instructional effectiveness and determine proper educational support for individual students. Collins and Wugalter (1992) pointed out that psychological research and theory is increasingly turning to longitudinal studies to monitor development over a period of time and test measurement methods using latent variables. The C-LCDM method is in its infancy. In fact, as of this study, there was only one reference. For longitudinal analysis, C-LCDM method should be further developed by incorporating Markov chain models, which are used to predict the probabilities of movement through stages over a specific time interval (transition probabilities), similar to latent transition analysis (Collins et al, 1994; Collins & Cliff, 1990; Martin et al., 1996; Velicer et al., 1996). Alternatively, we could gather multiple cross-sectional data in time and track changes of learning levels for descriptive purposes.

Second, the ultimate goal of applying C-LCDMs to the validation of learning progress models is to determine parameters in the C-LCDM model that can be generalized. In other words, based on the identified parameters, a student's stages might easily and quickly be estimated using their responses. That diagnostic algorithm could be embedded in an assessment technology.

Devising an assessment technology adapted for the complex and dynamic structure of mental models is essential because assessment is a fundamental step in feedback, revision, and reflection on learning (Pellegrino et al., 2001). That technology could enable a teacher to have a better sense of student learning and provide elaborate feedback and support. For example, McKeown (2009) used HIMATT with 40 actual classroom teachers. They managed to use that technology to diagnose student understanding and provide instructional support to different individuals even though the tool was new to them.

Third, this study discussed the assessment model and methods for diagnosing domain knowledge as an internally represented problem situation. Scholars generally accept that human cognition includes meta-cognition. Therefore, future studies should investigate the extent to which metacognition influences or interacts with the proposed stages of learning progress. Furthermore, meta-cognition might progress through qualitatively different stages, just as general cognition presumably does.

Considering that this study is only the first attempt to suggest a framework for measuring learning progress and to validate the model with data, more study is needed to conclude how many stages of knowledge structure might characterize learning progress. In addition, there are diverse theoretical accounts of the mental stages. Therefore, scholars should conduct further studies to test various conditions.

6.3 Closing Thoughts

Along with the theoretical models of learning progress, this study provides tools and methods for future research: (a) a set of parameters quantifying the attributes of knowledge structure; (b) a set of similarity measures applicable to the study of cognitive changes; and (c) a statistical approach (i.e., C-LCDM) to diagnosing the stages of learning progress. These models and methods for understanding levels of mastery of complex problems should help teachers develop learning environments that provide instructional support and feedback catered to the needs of individual learners. Thus, future studies should investigate instructional and feedback strategies associated with each developmental stage of learning progress. Instructional models based on diagnostic assessment could lead to the development of diverse instructional applications, such as intelligent tutoring systems.

References

Agresti, A. (2007). *An introduction to categorical data analysis*. Hoboken, NJ: Wiley Inter-Science.

Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 213–250). Greenwich, CT: JAI Press.

Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher, 32*(8), 10-14.

Alexander, P. A. (2004). A model of domain learning: Reinterpreting expertise as a multidimensional, multistage process. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 273-298). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Alexander, P. A., & Murphy, P. K. (1998). Profiling the differences in students' knowledge, interest, and strategic processing. *Journal of Educational Psychology, 90*, 435-447.

Alexander, P. A., Sperl, C. T., Buehl, M. M., Fives, H., & Chiu, S. (2004). Modeling Domain Learning: Profiles From the Field of Special Education. *Journal of Educational Psychology, 96*(3), 545–557. doi:10.1037/0022-0663.96.3.545

Anzai, Y., & Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction, 1*, 397-450.

Biggs, J., & Collis, K. (1982). *Evaluating the quality of learning: the SOLO taxonomy*. New York: Academic Press.

Bozard, J. L. (2010). *Invariance testing in diagnostic classification models*. Unpublished master's thesis, University of Georgia.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). Learning and transfer. In *How people learn: Brain, mind, experience, and school* (pp. 31–78). Washington, DC: National Academy Press.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology, 3*, 193-209.

Bransford, J. D., & Franks, J. J. (1972). The abstraction of linguistic ideas. *Cognitive Psychology*, 2, 331-350.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 717-726. doi:16/S0022-5371(72)80006-9

Carley, K., & Palmquist, M. (1992). Extracting, representing and analyzing mental models. *Social Forces, 70*, 215-225.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 1–58). New York: Academic Press.

Chen, Z., & Siegler, R. S. (2000). Intellectual development in childhood. In R.Sternberg (Ed.) *Handbook of intelligence* (pp. 92-116). New York: Cambridge.

Chi, M.T.H. (2006). Two approaches to the study of experts' characteristics. In K.A. Ericsson, N. Charness, P. Feltovich, & R. Hoffman (Eds.), *Cambridge Handbook of Expertise and Expert Performance* (pp. 121-130), Cambridge University Press.

Chi, M.T.H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 3-34). New York: Routledge.

Chi, M. T. H., & Glaser, R. (1985). Problem solving ability. In R.J. Sternberg (Ed.), *Human abilities: An information processing approach*. New York: W.H. Freeman.

Chi, M. T. H., Glaser, R., & Farr, M. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Choi, I., & Lee, K. (2009). Designing and Implementing a Case-Based Learning Environment for Enhancing Ill-Structured Problem Solving: Classroom Management Problems for Prospective Teachers. *Educational Technology Research and Development, 57*(1), 99–129. doi:10.2307/25619959

Clariana, R. B. (2010). Multi-decision Approaches for Eliciting Knowledge Structure. In *Computer-Based Diagnostics and Systematic Analysis of Knowledge* (pp. 41-59). Retrieved from http://dx.doi.org/10.1007/978-1-4419-5662-0_4

Collins, L. M. & Cliff, N. (1990). Using the longitudinal Guttmann simplex as a basis for measuring growth. *Psychological Bulletin, 108*, 128-134.

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

Collins L. M, Graham, J. W., Rousculp S. S., Fidler, P. L., Pan, J, & Hansen, W. B. (1994). Latent transition analysis and how it can address prevention research questions. *NIDA research monograph. 142*, 81-111.

Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research, 27*(1), 131-157.

Colman, A. M., & Shafir, E. (2008). Tversky, Amos. In N. Koertge (Ed*.), New dictionary of scientific biography* (Vol. 7, pp. 91-97). Farmington Hills, MI: Charles Scribner's Sons.

Coronges, K. A., Stacy, A. W., & Valente, T. W. (2007). Structural Comparison of Cognitive Associative Networks in Two Populations. *Journal of Applied Social Psychology*, *37*(9), 2097-2129. doi:10.1111/j.1559-1816.2007.00253.x

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

Curtis, M. B., & Davis, M. A. (2003). Assessing knowledge structure in accounting education: an application of Pathfinder Associative Networks. *Journal of Accounting Education*, *21*(3), 185-195. doi:10.1016/S0748-5751(03)00024-1

diSessa, A. A. (2006). A history of conceptual change research. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning science* (pp. 265-281). New York: Cambridge University Press.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction, 13*(5), 533-568. doi:10.1016/S0959-4752(02)00025-7

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: Free Press.

Ericsson, K. A. (2003). The acquisition of expert performance as problem solving: Construction and modification of mediating mechanisms through deliberate practice. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 31-83). New York: Cambridge University Press.

Ericsson, K. A. (2005). Recent advances in expertise research: A commentary on the contributions to the special issue. *Applied Cognitive Psychology, 19*, 233-241.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman, R. R. (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 685-706). Cambridge, UK: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245. doi:10.1037/0033-295X.102.2.211

Ericsson, K., & Simon, H. (May 1980). Verbal reports as data. *Psychological Review, 87*(3), 215–251. doi:10.1037/0033-295X.87.3.215

Ericsson, K., & Simon, H. (1993). *Protocol analysis: Verbal reports as data (2nd ed.)*. Boston: MIT Press. ISBN 0262050293.

Ericsson, K. A., & Staszewsli, J. J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 235–267). Hillsdale, NJ: Lawrence Erlbaum.

Eseryel, D., Ifenthaler, D., & Ge, X. (2013). Validation study of a method for assessing complex ill-structured problem solving by using causal representations. *Educational Technology Research and Development, 61*(3), 443–463. doi:10.1007/s11423-013-9297-2

Festinger, L. (1962). Cognitive dissonance. *Scientific American, 207*(4), 93-102.

Flavell, J. H. (1985). *Cognitive development* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Flavell, J. H., & Miller, P. H. (1998). Social cognition. In W. Damon & D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (pp. 851–898). New York: Wiley.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill.

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65*, 263–297.

Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of educational research*, *75*(1), 27-61.

Glaser, E. M., Abelson, H. H., & Garrison, K. N. (1983). *Putting knowledge to use: Facilitating the diffusion of knowledge and the implementation of planned change* (1st ed.). San Francisco, CA: Jossey-Bass.

Goldsmith, T. E., & Kraiger, K. (1997). Applications of structural knowledge assessment to training evaluation. In J. K. Ford (Ed.), I*mproving training effectiveness in organizations* (pp. 73–95). Mahwah, NJ: Lawrence Erlbaum Associates.

Goodman, C. M. (1987). The Delphi technique: a critique. *Journal of Advanced Nursing*, *12*(6), 729-734.

Greeno, J. G. (1989). Situations, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing* (pp. 285–318). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Grow-Maienza, J., Hahn, D., & Joo, C. (2001). Mathematics instruction in Korean primary schools: Structures, processes, and a linguistic analysis of questioning. *Journal of Educational Psychology, 93*(2), 363-376.

Guadagnoli, E., & Velicer, W. F., (1988). *Relation of sample size to the stability of component patterns. Psychological Bulletin, 103*(2), 265-275.

Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition, 50*(1-3), 171-188. doi:10.1016/0010-0277(94)90027-2

Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement.* New York, NY: Taylor & Francis Group.

Heinen, T. (1996). *Latent class and discrete latent trait models: Similarities and differences*. Thousand Oaks: Sage Publication.

Henson, R. A., Templin, J. L., & Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika,74* (2), 191-210.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15*, 332–340.

Hsu, C. & Sandford, B. (2007). The Delphi Technique: Making Sense of Consensus. *Practical Assessment Research & Evaluation, 12*(10). Available online: http://pareonline.net/getvn.asp?v=12&n=10

Ifenthaler, D. (2006). *Diagnose lernabhängiger Veränderung mentaler Modelle. Entwicklung der SMD-Technologie als methodologische Verfahren zur relationalen, strukturellen und semantischen Analyse individueller Modellkonstruktionen*. Freiburg: Universitäts-Dissertation.

Ifenthaler, D. (2010). Relational, structural, and semantic analysis of graphical representations and concept maps. *Educational Technology Research & Development, 58*(1), 81-97. doi:10.1007/s11423-008-9087-4.

Ifenthaler, D., Masduki, I., & Seel, N. M. (2009). The mystery of cognitive structure and how we can detect it: tracking the development of cognitive structures over time. *Instructional Science*. doi:10.1007/s11251-009-9097-6

Ifenthaler, D., & Seel, N. M. (2005). The measurement of change: Learning-dependent progression of mental models. *Technology, Instruction, Cognition, and Learning*, *2* (4), 321-340.

Johnson-Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.

Johnson-Laird, P.N. (2005a). Mental models and thoughts. In K. J. Holyoak (Ed.), *The Cambridge handbook of thinking and reasoning* (pp. 185-208). Cambridge University Press.

Johnson-Laird, P.N. (2005b). The history of mental models. In K. I. Manktelow, & M. C. Chung (Eds), *Psychology of reasoning: theoretical and historical perspectives* (pp. 179-212). Psychology Press.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45*(1), 65-94.

Jonassen, D. H, (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*(4), 63-85.

Jonassen, D. H. (2014). Assessing Problem Solving. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 269–288). Springer New York. Retrieved from ttp://link.springer.com/chapter/10.1007/978-1-4614-3185-5_22

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge. *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge.*

Kaplan, D. (2008). An overview of Markov chain methods for the study of stage-sequential development processes. *Developmental Psychology, 44*(2), 457-467.

Katz, J. J., & Postal, P. M.(1964). *An integrated theory of linguistic descriptions*. Cambridge: M.I.T. Press.

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

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development, 4*, 222-232.

Kim, M. (2012a). Cross-validation study on methods and technologies to assess mental models in a complex problem-solving situation. Computers in Human Behavior, 28 (2), 703-717. doi:10.1016/j.chb.2011.11.018

Kim, M. (2012b). Theoretically grounded guidelines for assessing learning progress: Cognitive changes in ill-structured complex problem-solving contexts. Educational Technology Research and Development, 60(4), 601-622. doi:10.1007/s11423-012-9247-4.

Kim, M. (2013). Concept map engineering: Methods and tools based on the semantic relation approach. Educational Technology Research and Development, 61(6), 951-978. doi: 10.1007/s11423-013-9316-3

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363-394. doi:10.1037/0033-295X.85.5.363

Lee, J., & Park, O. (2007). Adaptive instructional systems. In J. M. Spector, M. D. Merill, J. van Merrienboer & M. P. Driscoll (Eds.), *Handbook of research for educational communications and technology* (pp. 469-484). Routledge: Taylor & Francis Group.

Martin, R. A., Velicer, W. F., & Fava, J. L. (1996). Latent transition analysis to the stages of change for smoking cessation. *Additive Behaviors, 21*(1), 67-80.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*(1), 43-64.

McKeown, J. O. (2009). *Using annotated concept map assessments as predictors of performance and understanding of complex problems for teacher technology integration*. (Unpublished doctoral dissertation). Florida State University, Tallahassee, FL.

Monge, P. R., & Contractor, N. S. (2003). *Theories of communication networks*. New York: Oxford University Press.

Narayanan, V. K. (2005). Causal mapping: An historical overview. In V. K. Narayanan & D. J. Armstrong (Eds.), *Causal mapping for research in information technology* (pp.1-19). Hershey, PA: Idea Group Publishing.

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

Norman, D. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7-14). Hillsdale, NJ: Erlbaum.

Novak, J. D., & Cañas, A. J. (2006). *The origins of the concept mapping tool and the continuing evolution of the tool*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.106.3382

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & Management*, *42*(1), 15-29. doi:10.1016/j.im.2003.11.002

Opfer, J. E., & Siegler, R. S. (2004). Revisiting preschoolers' living things concept: A microgenetic analysis of conceptual change in basic biology. *Cognitive Psychology, 49*, 301-332.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.) (2001). *Knowing what students know*. Washinton, DC: National Academy Press.

Piaget, J. (1964). Development and learning. In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered* (pp. 7-20). Ithaca, NY: Cornell University.

Pirnay-Dummer, P. (2006). *Expertise und Modellbildung – MITOCAR*. Freiburg: Universitäts-Dissertation.

Pirnay-Dummer, P., & Ifenthaler, D. (2010). Automated knowledge visualization and assessment. In D. Ifenthaler, P. Pirnay-Dummer & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge*. New York: Springer.

Pirnay-Dummer, P., Ifenthaler, D., & Spector, J. (2010). Highly integrated model assessment technology and tools. *Educational Technology Research & Development, 58*(1), 3-18. doi:10.1007/s11423-009-9119-8.

Pretz, J. E., Naples, A. J., & Sternberg, R. J. (2003). Recognizing defining, and representing problems. In J. E. Davidson & R. J. Sternberg, *The psychology of problem solving* (pp. 3-30). New York: Cambridge University Press.

Rost, J., & Langeheine, R. (Eds.) (1997). *Applications of latent trait and latent class models in the social sciences*. New York: Waxmann.

Rupp, A. A., Sweet, S., & Choi, Y. (2010ba). *Modeling learning trajectories with epistemic network analysis: A simulation-based investigation of a novel analytic method for epistemic games*. Presented at the annual meeting of the International Society for Educational Data Mining (EDM), Pittsburgh, PA.

Rupp, A. A., Templin, J. L., & Henson, R. A. (2010b). *Diagnostic Measurement: Theory, Methods, and Applications*. New York: The Guilford Press.

Schlomske, N., & Pirnay-Dummer, P. (2008). Model based assessment of learning dependent change during a two semester class. In Kinshuk, Sampson, D., & Spector, M. (Eds.), *Proceedings of IADIS International Conference Cognition and Exploratory Learning in Digital Age 2008*, Freiburg, Germany, pp. 45-53.

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology, 9*, 1-16.

Schvaneveldt, R.W., Durso, F.T., Goldsmith, T.E., Breen, T.J., & Cooke, N.M. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies, 23,* 699-728.

Seel, N.M. (1999). Semiotics and structural learning theory. *Journal of Structural Learning and Intelligent Systems, 14* (1), 11–28.

Seel, N. M. (2001). Epistemology, situated cognition, and mental models: Like a bridge over troubled water. *Instructional Science, 29* (4-5), 403-427.

Seel, N. M. (2003). Model-centered learning and instruction. *Technology, Instruction, Cognition, and Learning, 1* (1), 59-85.

Seel, N. M. (2004). Model-centered learning environments: Theory, Instructional design, and ethics. In N. M. Seel & S. Dijkstra (Eds.), *Curriculum, plans, and processes in instruction design: International perspectives* (pp. 49-74). NJ: Lawrence Erlbaum Associates, Inc.

Seel, N. M., & Dinter, F. R. (1995). Instruction and mental model progression: Learner-dependent effects of teaching strategies on knowledge acquisition and analogical transfer. *Educational Research and Evaluation, 1*(1), 4-35.

Segers, M. (1997). An alternative for assessing problem-solving skills: The overall test. *Studies in Educational Evaluation, 23*(4), 373-398.

Shin, N., Jonassen, D. H., & MaGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science Teaching, 40*(1), 7- 27.

Siegler, R. S. (2005). Children's Learning. *American Psychologist, 60*, 769-778.

Siegler, R. S., Thompson, C. A., & Opfer, J. E. (2009). The logarithmic-to-linear shift: One learning sequence, many tasks, many time scales. *Mind, Brain, and Education, 3*, 143-150.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61*, 394-403.

Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitive Psychology, 8*, 165–190.

Shute, V. J., & Zapata-Rivera, D. (2008). Using an evidence-based approach to assess mental models. In D. Ifenthaler, P. Pirnay-Dummer, J. M., Spector (Eds.), *Understanding Models for Learning and Instruction: essays in honor of Norbert M. Seel* (pp. 23-41). New York: Springer.

Sinnott, J. D. (1989). A model for solution of ill-structured problems: Implications for everyday and abstract problem solving. In J. D. Sinnott (Ed.), *Everyday problem solving: Theory and applications* (pp. 72-99). New York: Praeger.

Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences, 3*(2), 115-163.

Snow, R.E. (1990). New approaches to cognitive and conative assessment in education. *International Journal of Educational Research, 14* (5), 455–473.

Stigler, J. W., & Stevenson, H. W. (1991). How Asian teachers polish each lesson to perfection. *American Educator, 15*, 12-20.

Spector, J. M. (2004). Problems with problem-based learning: Comments on model-centered learning and instruction in Seel (2003). *Technology, Instruction, Cognition and Learning, 1*(4), 359–374.

Spector, J. M. (2008). Expertise and dynamic tasks. In H. Qudrat-Allah, J. M. Spector, & P. I. Davidsen (Eds.), *Complex decision making: Theory and practice*. Berlin: Springer-Verlag.

Spector, J. M., & Koszalka, T. A. (2004). *The DEEP methodology for assessing learning in complex domains* (Final report to the National Science Foundation Evaluative Research and Evaluation Capacity Building). Syracuse, NY: Syracuse University.

Spiro, R. J., Feltovich, P. J., & Coulson, R. L. (1996). Two epistemic world-views: Prefigurative schemas and learning in complex domains. *Applied Cognitive Psychology, 10*, s51-s61.

Steven, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NJ: Erlbaum.

Taricani, E. M., & Clariana, R. B. (2006). A Technique for Automatically Scoring Open-Ended Concept Maps. *Educational Technology Research & Development*, *54*(1), 65-82. doi:Article

Templin, J. L. (2004). *Generalized linear mixed proficiency models*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327–352.

Tversky, A., & Shafir, E. (2004). *Preference, belief, and similarity: selected writings*. London, England: MIT Press.

Velicer, W. F., Martin, R. A., & Collins, L. M. (1996). Latent transition analysis for longitudinal data. *Addition, 91* (Supplement), s197-s209.

Vosniadou, S. (2003). Exploring the relationships between conceptual change and intentional learning. In G. M. Sinatra & P. R. Pintrich (Ed.), *Intentional conceptual change* (pp. 377-406). Mahwah, NJ: Erlbaum.

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 3-34). New York: Routledge.

Vygotsky, L. (1934/1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications*. Cambridge University Press.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D. B Harris (Ed.), *The concept of development: An issue in the study of human behavior* (pp. 125–148). Minneapolis, MN: University of Minnesota.

Wood, P. K. (1983). Inquiring systems and problem structures: Implications for cognitive developments. *Human Development, 26*, 249-265.

Zsambok, C. E., Klein, G. (Eds.). (1997). *Naturalistic decision making*. Mahwah, NJ: Lawrence Erlbaum Associates.

Appendix A

Case Study

Directions: Read the case study described below and then prepare a response to the questions below (**written response with at least 350 words is required for each question**):

Assume that you have been involved in evaluating a media implementation project in an urban inner middle school. At the beginning of the school year, all of the students assigned to four subject area teachers (math, language arts, social studies and science) in the seventh grade at the middle school were given tablet PCs (laptop computers also equipped with a stylus/pen and a touchscreen that can be written upon) and were also given wireless internet access at home and at school for a entire year.

The students took the tablet PCs home every evening and brought them to class every day. The teachers were also provided with tablet PCs 24/7 (24 hours a day, every day of the week) for the entire year. The teachers and students were trained on how to use the tablet PCs. Moreover, all of the curriculum materials (textbooks, workbooks, student study guides, teacher curriculum guides, some activities, tests, etc.) were installed on the tablet PCs or were accessible through the tablet PCs.

Your job as one of the evaluators for the project was to examine how this innovation (providing teachers and students with tablet PCs 24/7) changed the way instruction was presented in the classrooms of the four teachers. Results indicated that the innovation had very little effect on the manner in which instruction took place in the teachers' classrooms.

1. Based on what you have learned about the use of technology in education, describe what concepts, issues, factors, and variables are likely to have contributed to the fact that the introduction of the tablet PCs had very little effect on the instructional practices that were employed in the classes.

2. Describe the strategies that could have been employed to help mitigate the factors that you think contributed to the minimal effect the tablet PCs had on instructional practices. When you answering this question, use the concepts, factors, and variables you described in the question 1 or add other assumptions and information that would be required to solve this problem.

Appendix B

Reference model to the Question 1

Technology implementations usually begin with an identified instructional need. *Instructional need was likely not fully identified due to* insufficient study of how instructional practices in the classroom were being conducted already without the technology. One big issue is defining what a successful integration or change in instructional practice actually is. While *teachers* in the situation may have felt that they knew this already, the assumptions inherent in a design situation need to be articulated and checked if the assumptions are not to distort the design space by which *instructional practices are manipulated*. Teachers didn't have enough professional development using the technology in classroom teaching and learning, on ways to integrate use into their teaching, and best practices with regard to effective educational use. Teacher professional development that discusses not just technical know-how but also pedagogy could help teachers realize how to do things differently that takes full advantage of the affordances of the tablets. Training as a professional development should be extensive including teacher beliefs and attitude. Teacher beliefs play a role in adopting new practices and changing their instructional practice. Teachers may not believe that students learn with laptops, and thus do not use *laptops* in their instruction. The only support teachers had during implementation was technical support; Teachers lacked a mentor who could assist them as instructional issues arose throughout the year. Mentoring on additional and advanced uses of the technology in the classroom is critical for teachers to increase their skills and maintain their motivation in utilizing the technology. In addition, mentor could help teachers to maintain the belief that these efforts will have positive results. There are concerns that the environment does not support change. An ongoing supportive environment where teachers initially learn how to use the technology, how to use the technology with their content, and how to continue to develop their expertise in the technology and incorporating it to the classroom is critical. Environment could include a culture that does not support the desired performance.
For example, the lack of incentives to make effective use of a new technology could also contribute to lack of use. The intervention seems to have been applied to this community rather than involving *teachers* from the beginning as collaborators in its design and modification. Teachers were not involved in the decision to implement the new media; thus, they did not fully "buy into" the plan.

Appendix C

Cross-tabulation of School Year and Classes

	Class					
School Year	Acclimation	Competence	Proficiency	Total	$x^2(df)$	p
Freshman	3	4	2	9	1.914 (6)	.927
Sophomore	8	20	9	37		
Junior	17	23	12	52		
Senior	10	15	10	35		
Total	38	62	33	133		

Note. $^*p \leq .05$.

Cross-tabulation of Interest and Classes

Interest in becoming	Class					
a teacher	Acclimation	Competence	Proficiency	Total	$x^2(df)$	p
Not interested at all	5	10	2	17	2.813 (8)	.946
Not so interested	9	14	9	32		
Somewhat Interested	9	14	9	32		
Interested	5	8	6	19		
Very interested	10	16	7	33		
Total	38	62	33	133		

Note. $^*p \leq .05$.

Cross-tabulation of Teaching Experience and Classes

	Class					
Teaching Experience	Acclimation	Competence	Proficiency	Total	$x^2(df)$	p
None	33	59	32	9	9.032 (2)	.011
Yes but less than 2 years	5	1	0	37		
Total	38	60	32	133		

Note. $^*p \leq .05$.

Cross-tabulation of Prior Coursework and Classes

	Class					
Prior Coursework	Acclimation	Competence	Proficiency	Total	$x^2(df)$	p
None	20	22	12	54	3.224 (4)	.521
Other education courses	16	36	19	71		
Instructional Design courses	2	4	2	8		
Total	38	62	33	133		

Note. $^*p \leq .05$.