Journal of the Association for Information Systems

Research Paper ISSN: 1536-9323

Understanding the Role of IS and Application Domain Knowledge on Conceptual Schema Problem Solving: A Verbal Protocol Study

Vijay Khatri
Indiana University
vkhatri@indiana.edu

Iris Vessey
The University of Queensland
i.vessey@business.uq.edu.au

Abstract:

One of the most neglected areas of information systems research is the role of the domain to which researchers apply IS methods, tools, and techniques; that is, the application domain. For example, little prior information systems (IS) or related research has examined how IS and application domain knowledge (ISDK and ADK, respectively) influence how individuals solve conceptual schema problem-solving tasks. In this research, we investigate the effects of both ISDK and ADK on two types of conceptual schema problem-solving tasks: schema based and inferential. We used verbal protocol analysis to explore the roles that ISDK and ADK play in the problem-solving processes participants use when addressing these tasks. We found that, for the two types of conceptual schema problem-solving tasks, ADK and ISDK have similar effects on problem-solving processes. That is, we found that, for schema-based problem-solving tasks, participants used focused (depth-first) processes when the application domain was familiar as did participants with greater IS domain knowledge. We also found that, for inferential problem-solving tasks, participants used exploratory (breadth-first) processes when the application domain was familiar as did participants with greater IS domain knowledge. We then show how cognitive psychology literature on problem solving can help explain the effects of ISDK and ADK and, thus, provide the theoretical foundation for analyzing the roles of each type of knowledge in the process of IS problem solving.

Keywords: Conceptual Schema Understanding, Schema-based Problem-solving Tasks, Inferential Problem-solving Tasks, Protocol Analysis, Problem-solving Processes.

Maung Sein was the accepting senior editor. This article was submitted on May 24, 2015 and went through two revisions.

1 Introduction

Researchers have long acknowledged that information systems (IS) problem solving, "from concept definition to product retirement", involves application domain problem solving using a software solution (Blum, 1989, p. 502). This perspective highlights the role that both the IS and application domains play in IS problem solving. By application domain, we mean the area of business that IS problem solving addresses (e.g., accounting, production, sales). By IS domain, we mean areas in IS that have significant amounts of specialized knowledge (e.g., conceptual modeling, software development, software modification). Hence, the technical, IS domain includes representations, methods, techniques, and tools that form the basis for the initial and on-going development of application systems—systems that are applied to a specific area of business (i.e., the application domain). Therefore, knowledge of both the IS and application domains (ISDK and ADK, respectively) are important in IS problem solving.

The majority of prior research that has examined the role of the application domain in IS problem solving has examined software-related tasks (Hadar, Soffer, & Kenzi, 2014; Purao, Russi, & Bush, 2002; Shaft & Vessey, 1995, 1998, 2006; Vessey & Conger, 1993). The role of application domain knowledge in data-related tasks has received considerably less attention. Burton-Jones and Weber (1999) and Khatri, Vessey, Ramesh, Clay, and Park (2006), for example, focus on the role of the application domain in conceptual data modeling.

Although conceptual modeling of data-related requirements is a small phase in the overall organization and use of data, it has a greater impact on software development than any other phase (Witt & Simsion 2004). Further, data and, therefore, conceptual data models have gained increasing recognition as corporate assets (Khatri & Brown, 2010). For example, conceptual modeling has a significant role to play in meeting regulatory reporting requirements (Delbaere & Ferreira, 2007). Conceptual modeling is also central to master data management (Power 2011). Further, conceptual modeling can influence data quality (see, e.g., Wang & Strong, 1996).

Conceptual modeling involves developing and using conceptual schemas, abstract representations of the structure of the data relevant to a specific area of application (Hoffer, Prescott, McFadden, 2007). Understanding conceptual schemas is particularly important given that software maintenance accounts for 50-75 percent of software lifecycle costs (Edberg, Ivanova, & Kuechler, 2012; Koskinen, 2010), an estimated 50-90 percent of which is spent on understanding software and design diagrams (Robson, Bennett, Cornelius, & Munro, 1991). Given the importance of interacting with conceptual schemas, prior research has examined extensively conceptual schema understanding using different types of tasks (see, e.g., Burton-Jones & Meso, 2006; Lukyanenko, Parsons, & Wiersma, 2015; Shanks, Tansley, Nuredini, Tobin, & Weber, 2008)—those that employ working memory alone or those that integrate information obtained from the schema with that in long-term memory. Such tasks involve both *surface-level* and *deep-level understanding* of the "real world" (i.e., the application domain) (Saghafi & Wand, 2014).

While researchers have often acknowledged the role of the application, or real-world, domain (see, e.g., Burton-Jones & Meso, 2008), they have rarely examined it. As far as we are aware, only three prior studies have examined explicitly the role of the application domain in conceptual modeling. First, Burton-Jones and Weber (1999) found that ADK improved performance on tasks that required deep-level understanding when the participants were required to solve the problem based on ontologically unsound representations of relationships with attributes. Second, Khatri et al. (2006) found that ADK was effective in helping individuals solve conceptual schema understanding tasks when the tasks required deep-level understanding but not when they required surface-level understanding. Third, Bera, Burton-Jones, and Wand (2009) found that, for tasks that required deep-level understanding, role guidelines helped individuals when they were moderately familiar or unfamiliar with the application domain; the guidelines did not help, however, when the individuals were quite familiar with the application domain.

In this paper, we build on prior studies that suggest that ADK aids deep-level understanding to determine how ADK affects problem-solving processes on *conceptual schema problem-solving tasks* (Saghafi & Wand, 2014), which may be either schema based (see, e.g., Khatri et al., 2006; Shanks, Tansley, Nuredini, Tobin, & Weber, 2002) or inferential (see, e.g., Burton-Jones & Meso, 2006; Gemino & Wand, 2005) in nature. We conducted an in-depth exploratory study to examine the effect of ISDK and ADK on problem solving based on these two types of tasks. Specifically, we address the following research question:

> **RQ**: How do IS and application domain knowledge influence the way in which individuals solve conceptual schema problem-solving tasks?

To examine our research question, which focuses on the process of deep-level schema understanding, we examine search for information cues in the conceptual schema and/or the data dictionary. The Humans as Information Processing Systems (HIPS) paradigm provides a well-established theoretical basis for the fact that heuristic search is the process by which intelligence arises (Newell & Simon, 1976). Researchers in cognitive science, therefore, typically conceive of problem solving in terms of search. This notion is supported by a significant body of cognitive research in IS (see, e.g., Browne & Parsons 2012; Davern et al., 2012). As such, we focus our study on the roles of ISDK and ADK in information processing—acquisition and use—when participants solve conceptual schema problem-solving tasks.

In an exploratory study, we examined individuals' search processes by conducting in-depth analyses of verbal protocol data collected during a think-aloud problem-solving process (Newell & Simon 1972). In an idiographic analysis, we examined individual problem-solving processes visually using transition graphs. We followed the idiographic analysis with a nomothetic analysis by conducting an aggregate analysis of all participants' processes. We discovered that ADK and ISDK had similar effects on problem-solving processes for both types of conceptual schema problem-solving tasks. That is, we found that participants familiar with the application domain and those with greater IS domain knowledge used focused (depth-first) processes. We also found that, for inferential problem-solving tasks, participants familiar with the application domain and those with greater IS domain knowledge used exploratory (breadth-first) processes (Newell & Simon, 1972).

We contribute to existing knowledge by: 1) characterizing the problem-solving processes essential to solving each type of task, 2) demonstrating the importance of ISDK and ADK to solving both types of tasks, 3) identifying the fact that participants manifest limitations in cognitive capacity when they lack either ISDK or ADK, 4) presenting theory that explains the roles of ISDK and ADK in solving conceptual schema understanding tasks generally and conceptual schema problem-solving tasks in particular. Our examination of IS problem-solving processes, therefore, provides insights into "how" ISDK and ADK influence problem solving.

The paper proceeds as follows. In Section 2, we present the roles of ISDK and ADK in conceptual schemas and in solving schema-based and inferential problem-solving tasks. We also present the cognitive underpinnings of the approach we used to investigate our research question. In Section 3, we present our research methodology. In Section 4, we present our analysis and findings. In Section 5, we interpret our research based on the cognitive psychology literature on problem solving and state relevant propositions. We also present the contributions our research makes to knowledge in the area of conceptual modeling and discuss the implications our research for future research, for teaching, and for practice. Finally, we present our conclusions in Section 6.

2 Foundations

In this section, we discuss the foundation we used to examine the processes that individuals use to solve problems based on conceptual schemas. We addressed first the roles played by ISDK and ADK in conceptual schema problem solving and second the types of tasks we examine. We then investigate the nature of problem solving as a forerunner to establishing the way in which we investigate problem-solving processes.

2.1 Conceptual Schemas

We first examine the ways in which conceptual schemas represent ISDK and ADK. Second, we examine the types of conceptual schema problem-solving tasks that we addressed in our study.

2.1.1 Roles of ISDK and ADK in Conceptual Schemas

From the viewpoint of ISDK, considerable formalization of the characteristics of data has taken place over the past four decades (see, among others, Chen, 1976; Codd, 1970; Elmasri & Navathe, 2006). Formalization is reflected in the representation of the semantics of entity types and attributes, as well as numerous types of relationships (see Chen, 1976; Elmasri & Navathe, 2006). While researchers may need to further clarify nuances in the representation of data semantics (see, e.g., Sheth 1995), practitioners model unambiguously a large volume of structured organizational data using current data modeling formalisms (Geiger 2010). In this research, we focus on such unambiguous representations of data semantics and the associated understanding of schema semantics, which we characterize as a well-structured problem area. Furthermore, we do so in terms of the most widely used conceptual modeling formalism (Davies, Green, Rosemann, Indulska, & Gallo, 2006; Fettke, 2009), the ER model (Chen, 1976).

From the viewpoint of ADK, researchers have used a conceptual modeling formalism to develop conceptual schema— an abstract representation of the structure and interrelationships among the data of interest in an organization (i.e., the entities, events, activities, and their associations) (Hoffer et al., 2007). These entity types, attributes, and relationships represent ADK; that is, a conceptual schema highlights the characteristics of data in a given application domain and the way in which the data are interrelated. Figure A1 presents a conceptual schema for a sales domain. It includes several entity types (e.g., PRODUCT, PRODUCT LINE, and SALES PERSON, each with several attributes) and several relationships (e.g., the relationship exists in between PRODUCT and PRODUCT LINE). Entity types and relationships that link them, represented as icons (rectangles and diamonds, respectively) in the graphical conceptual schema, refer to ISDK, while the textual terms PRODUCT, PRODUCT LINE, and SALES PERSON represented in the icons designate the business area to which a conceptual modeler applies the ISDK. The relationship between PRODUCT and PRODUCT LINE is a mandatory-many (1:1 and 1:M) relationship (ISDK) that indicates that a PRODUCT can exist in exactly one PRODUCT LINE and that a PRODUCT LINE may have many associated PRODUCTs.

In summary, a conceptual schema is a problem representation that a conceptual modeler uses as a basis for problem solving. By representing data structures and interrelationships in the application domain of interest, a conceptual schema uses the notations of the ER model (i.e., the syntax) to represent the application's semantics. Given that the ER model's notations "stand for" concepts in the real world, a conceptual schema represents ISDK explicitly and ADK implicitly. That is, a conceptual schema foregrounds ISDK.

2.1.2 Conceptual Schema Problem-solving Tasks

As we mention above, we focus on conceptual schema problem-solving tasks that require deep-level understanding (Saghafi & Wand, 2014). For such tasks, ADK plays a role when individuals solve both schema-based and inferential problem-solving tasks. We consider each in turn.

Schema-based problem-solving tasks (see, e.g., Khatri et al., 2006; Shanks, Nuredini, Tobin, Moody, & Weber, 2003) resemble query tasks: respondents need to determine whether, and how, a schema makes available certain information they need to address a task. Given that the schema makes available the desired information, the respondent may solve such a task using the schema alone. The following extract exemplifies this type of problem-solving task:

> *Managers in the finance and marketing divisions need to decide which products to keep in their product portfolio. These decisions are based on measures of advertising budget, miscellaneous expenditure, and target audience for a given product line. Based on the material provided, can you find an answer to the above problem? If yes, describe how you would find the answer. Please be specific.* (Khatri et al., 2006, p. 97)

Inferential problem-solving tasks (see, e.g., Bodart et al., 2001; Burton-Jones & Weber, 1999; Gemino, 1999; Gemino & Wand, 2003; Shanks et al., 2003; Shanks, Nuredini, & Weber, 2004; Shanks et al., 2002) are the most complex of the schema-understanding tasks identified to date because they require inferential reconstruction (Gemino & Wand, 2003). Individuals need to infer what is plausible in the context of the semantics that a schema represents (Bodart et al., 2001). Given that the problem statement does not contain all the information they need to solve an inferential problem-solving task, respondents do not know what actions they need to take (Chi & Glaser, 1984). An example of an inferential problem-solving task is: "Some manufacturers do not manufacture certain products. Write as many reasons as you can think of that may have led to this situation." (see, e.g., Bodart et al., 2001; Gemino, 1999).

2.2 Investigating Processes in Solving Conceptual Schema Problem-solving Tasks

In this section, we present our approach to examining problem-solving processes that individuals use in conceptual schema problem solving. We focus on the nature of problem solving in general and on solving problems based on graphical problem representations in particular.

The cognitive psychology community has examined in depth problem-solving processes over an extended period—an examination that researchers have often framed in terms of domain knowledge. Studies in a variety of domains (e.g., chess, medical diagnosis, musical performance, programming, and software domains) have revealed that individuals use a remarkably similar set of problem-solving skills based on search processes (see, e.g., Chi, Feltovich, & Glaser, 1981; Mayer, 1997).

Search strategies are particularly salient when problem representations are graphical in nature. Indeed, Petre (1995, p. 37) notes that "Effective search strategies are particularly important with graphics since there are no cues to navigation in most diagrams and, therefore, the reader of graphical notation must first identify an appropriate inspection strategy”. Similarly, in terms of understanding conceptual schemas, search strategies provide valuable insights into visual processing based on graphical conceptual models for which there are few or no cues to aid navigation (Hungerford, Hevner, & Collins, 2004; see also Browne & Parsons 2012; Davern et al., 2012). Therefore, search strategies based on information cues in the schema and / or the data dictionary appear to be particularly relevant when solving problems based on conceptual schema. We expect that the more structured the task, the more focused the search strategy and, therefore, the fewer the number of information cues needed to solve the problem. Given that schema-based problem-solving tasks are more structured than inferential problem-solving tasks, we expect that problem solvers will display more-focused search processes (referred to as “depth-first”) when addressing schema-based problem-solving tasks and more exploratory search processes (referred to as “breadth-first”) when addressing inferential problem-solving tasks. Likewise, we expect that more experienced problem solvers will use more-focused problem-solving processes than less experienced problem solvers. For further information on breadth-first and depth-first problem-solving processes, see Adelson and Soloway (1985), Greeno (1978), Jeffries, Turner, Poison, and Atwood (1981), Johnson et al. (1981), Malhotra, Thomas, Carrol, and Miller (1980), and Rist (1989).

We evaluate the effect of ISDK by examining the type of processes employed by two groups of participants who differed on the extent of their ISDK. We refer to the two groups as having high and low ISDK. To evaluate the effect of ADK, we examined the type of problem-solving processes employed in two application domains (one the participants were familiar with and one they were not). We refer to these application domains as familiar and unfamiliar application domains, respectively. We refer to the level of knowledge in the application domain in terms of being familiar or unfamiliar: in familiar domains, ADK is high; in unfamiliar domains, ADK is low.

3 Research Methodology

To answer our research question (i.e., How do IS and application domain knowledge influence the way in which individuals solve conceptual schema problem-solving tasks?), we employed verbal protocol analyses. Problem solvers with different levels of IS domain knowledge responded to two problem-solving tasks of different levels of structure in familiar and unfamiliar application domains.

3.1 Task Setting

We investigated sales and hydrology as our two application domains. We expected that participants drawn from a business school (see Section 3.2) would be more familiar with a sales application and less familiar with a hydrology application. Because the ER model is the most commonly used conceptual modeling formalism (42% of respondents indicated frequent use, which exceeded the number of respondents (21%) who frequently used UML’s class diagrams) (Davies, Green, Rosemann, & Gallo, 2004), we investigated our research question based on common aspects of the ER and EER conceptual models (see Chen, 1976; Elmasri & Navathe, 2006, respectively).

3.2 Participants

Twelve undergraduate students from two sections of a data management course offered in the business school of a large U.S. university and who were well versed in conceptual modeling took part in the study (see our analysis of sample sizes in studies such as ours in Section 6). Participation in the study was voluntary, and we gave participants US$30 to complete the experiment. Appendix B shows participants’ demographics. They were between 20 and 25 years and had a high-school diploma, some work experience, and little database-related work experience. Given the characteristics of our respondents, we generalize our results to new recruits entering the role of data analysts in practice.

3.3 Experimental Design

Participants who demonstrated high and low ISDK each completed eight tasks, two tasks in each of the two types of problem-solving categories examined in both familiar and unfamiliar application domains. We randomly assigned participants to groups using either the ER or EER models. The problem-solving tasks investigated in this research involved only entity types/relationships and attributes (henceforth, ERA); that is, concepts common across ER and EER models. Further, we counterbalanced the presentation sequence of the two schemas (familiar and unfamiliar) to control for any order effects.

3.3.1 Operationalizing IS Domain Knowledge

To investigate the influence of IS domain knowledge, we formed groups of participants with high and low ISDK based on their scores on tasks researchers have frequently used to gauge schema comprehension (Khatri et al. 2006); that is, syntactic and semantic comprehension tasks (see Table 1). These two types of schema-comprehension tasks (see also Appendix C) focus on the notations foregrounded in the schema. One can, therefore, solve them by referencing the conceptual schema directly using ISDK alone (Khatri et al. 2006). Such scores are an appropriate measure of IS knowledge because the cognitive psychology literature has established that knowledge of surface features forms the foundation for developing higher forms of knowledge (Anderson, 1982, 1996). Based on the syntactic performance in Table 1, we note that all of our participants understood the notations of the ER model. On the other hand, our participants' semantic performance varied from a high of 39 to a low of 24 (out of a maximum score of 40).

Table 1. Performance of Participants on Syntactic and Semantic Comprehension Tasks

Participant	Syntactic performance[1] (maximum: 20)	Semantic performance[1] (maximum: 40)	Total (maximum: 60)
H-1	20	39	59
H-2	20	38	58
H-3	20	38	58
H-4	20	37	57
H-5	20	36	56
H-6	20	35	55
		Mean	**57.17**
L-1	20	31	51
L-2	19	31	50
L-3	20	30	50
L-4	18	28	46
L-5	19	28	47
L-6	18	24	42
		Mean	**47.67**

We ranked participants based on their "total" scores (see Table 1) and then selected the six highest performers to form the group with high IS domain knowledge (H-ISDK; labeled H-x in Table 1),and the six lowest performers to form the group with the low IS domain knowledge (L-ISDK; labeled L-x). Out of a maximum score of 60, the six highest and the six lowest performers scored an average of 57.17 and 47.67, respectively. The significant gap in performance figures between the high- and low-performing participants also supports our approach.

3.3.2 Operationalizing Application Domain Knowledge

As we show above, our experimental design called for using two domains with which our participants would not be equally familiar. We refer to these application domains as familiar (sales) and unfamiliar (hydrology). As a manipulation check on ADK prior to the experiment proper, we examined participants' knowledge of terms that mapped to concepts on the conceptual schemas with which they later interacted, five in each of the sales (product line, sales person, warehouse, area headquarter and manufacturer) and hydrology (seep, playa, bore hole, lithology and pump) domains. Hence, this exercise highlighted what participants knew about aspects of each domain. We then asked the participants to rate their familiarity with sales and hydrology applications on a seven-point Likert scale (where 7 = high and 1 = low familiarity) (see Table 2).

[1] Syntactic performance in sales and hydrology.

Table 2 shows that the self-reported familiarity of the participants was far greater in the sales (5.00) than in the hydrology (1.33) domain.

Table 2. Participants' Knowledge of the Application Domain

Participants	Sales	Hydrology
H-1	5	1
H-2	6	1
H-3	5	1
H-4	6	1
H-5	5	2
H-6	5	1
L-1	3	1
L-2	5	2
L-3	5	1
L-4	5	3
L-5	5	1
L-6	5	2
Mean	**5.00**	**1.33**

3.4 Experimental Materials

We presented each participant with two schemas (one in the familiar domain (sales) and the other in the unfamiliar domain (hydrology)) and their corresponding data dictionaries. The data dictionaries included application-oriented descriptions of each entity type/relationship and attribute on the schema. The schemas were syntactically equivalent; only the labels used for entity types, relationships, and attributes differed. We adapted the schema and the data dictionaries from Khatri et al. (2006). Appendix A presents the schema and corresponding data dictionaries. The sales schema (Figure A1) was a typical order-processing application that included concepts such as SALES AREA, SALES TERRITORY, PRODUCT, PRODUCT LINE, and MANAGER. Table A1 presents an excerpt from the corresponding data dictionary. We adapted the hydrology schema (Figure A2) from a schema for a ground water application at the U.S. Geological Survey. This application included hydrological concepts such as SEEP, PLAYA, BORE HOLE, CASING, and ACCESS TUBE. Table A2 presents an excerpt from the corresponding data dictionary.

Our participants responded to tasks that focused only on ERA. The tasks were structurally equivalent in each domain; that is, participants needed to address structurally corresponding entity types, relationships, and attributes to respond to the corresponding task in each application domain. In Sections 3.4.1 and 3.4.2, we first present the tasks that we used in this study. We then describe the information cues that respondents needed to solve these tasks.

3.4.1 Schema-based Problem-solving Tasks

Our participants responded to two schema-based problem-solving tasks in each of the sales and hydrology domains (see Table 3). We refer to the two schema-based problem-solving tasks as Tasks 1a and 1b.

In the familiar application domain, solving Task 1a (see Table 3) requires making a decision on retaining products in a PRODUCT LINE according to a set of criteria. The schema presents the criteria as attributes of PRODUCT LINE. A respondent does not need to analyze entity types other than PRODUCT LINE. The corresponding entity type in the unfamiliar application domain is BORE HOLE SITE.

Table 3. Schema-based Problem-solving Tasks

Task 1a (Sales)	Managers in the finance and marketing divisions need to decide which products to keep in their product portfolio. These decisions are based on measures of advertising budget, miscellaneous expenditure, and target audience for a given product line. Based on the material provided, can you find an answer to the above problem? If yes, describe how you would find the answer. Please be specific.

Table 3. Schema-based Problem-solving Tasks

Task 1a (Hydrology)	Geologists and hydrologists need to decide which bore holes to include in their groundwater study. These decisions are based on measures of leakance, horizontal conductivity, and vertical conductivity at a bore hole site. Based on the material provided, can you find an answer to the above problem? If yes, describe how you would find the answer. Please be specific.
Task 1b (Sales)	A group of customer service managers needs to understand the defects associated with recently delivered products. We need to find the address and CEO of a manufacturer that produced products in the last five years. Based on the material provided, can you find an answer to the above problem? If yes, describe how you would find the answer. Please be specific.
Task 1b (Hydrology)	A group of earth scientists needs to understand the rock formations associated with recently constructed bore holes. We need to find the age and formation name of a lithology that is related to bore holes constructed in the last five years. Based on the material provided, can you find an answer to the above problem? If yes, describe how you would find the answer. Please be specific.

In the familiar application domain, solving Task 1b requires finding the CEO and address of a MANUFACTURER that produced PRODUCTs in the last five years. A response to this task requires participants to determine the history of the PRODUCTs that are associated with MANUFACTURER. The schema does not, however, present the dates on which the manufacturer produced the PRODUCTs. Therefore, a respondent cannot answer this question. The respondent does not need to analyze entity types other than PRODUCT and MANUFACTURER. In the unfamiliar application domain, the corresponding entity types are LITHOLOGY and the associated BORE HOLE history.

We refer to the above solutions as parsimonious. Appendix D shows the transition graphs for the most parsimonious solutions of Tasks 1a and 1b.

3.4.2 Inferential Problem-Solving Tasks

Our participants responded to two inferential problem-solving tasks in each of the sales and hydrology domains (see Table 4). We refer to the two inferential problem-solving tasks as Tasks 2a and 2b.

Table 4. Inferential Problem-Solving Tasks

Task 2a (Sales)	Some manufacturers do not manufacture certain products. Write as many reasons as you can think of that may have led to this situation.
Task 2a (Hydrology)	Some lithologies are not associated with any bore hole. Write as many reasons as you can think of that may have led to this situation.
Task 2b (Sales)	A sales person needs to manage a sales area. Write as many factors as you can think of that they need to consider.
Task 2b (Hydrology)	A source agency is planning to observe a spring site. Write as many factors as you can think of that they need to consider.

In the familiar application domain, the solution of Task 2a (see Table 4) requires exploring the relationship between MANUFACTURER and PRODUCT (e.g., patent expiry date for a given product by a manufacturer), and perhaps also between PRODUCT and ORDER and between PRODUCT and PRODUCT LINE. In the unfamiliar application domain, a similar situation arises, this time between CONSTRUCTION AGENCY and BORE HOLE and perhaps between BORE HOLE and LITHOLOGY and between BORE HOLE and BOREHOLE SITE.

In the familiar application domain, solving Task 2b requires exploring the relationships between SALES PERSON and SALES AREA (e.g., population of the sales areas that the sales person is supposed to manage, the assigned budget, etc.), SALES PERSON and PRODUCT LINE, SALES AREA and SALES TERRITORY, and even perhaps PRODUCT LINE and WAREHOUSE as well as SALES TERRITORY and AREA HEADQUARTER. In the unfamiliar application domain, the corresponding entity types are SOURCE AGENCY, SPRING SITE, BOREHOLE SITE, PLAYA, ACCESS TUBE, and SEEP.

3.5 Pilot Study and Experimental Procedure

To test that the procedures were effective, we conducted a pilot study with students who had conceptual modeling experience. The pilot study helped us to eliminate ambiguity in question wording, test the experimental procedures, and determine the length of time that the experiment would take to complete. We used the pilot data to develop the coding scheme for the experiment proper.

In the study itself, we used a common script to conduct 12 individual sessions. Following an introduction to the study, participants completed a background questionnaire that sought demographic information and measured their a priori familiarity with the sales and hydrology application domains. We then gave participants an information sheet that described the syntax of the assigned model and had them view a PowerPoint video (developed using Camtasia Studio) that recapped key conceptual modeling concepts. Next, we instructed participants on the verbal protocol (think-aloud) technique that we used to collect experimental data and had them complete a practice exercise to become familiar with thinking aloud while problem solving. Participants became conversant with the experimental schema by responding to syntactic and semantic conceptual schema-comprehension tasks. Finally, participants completed two sets of experimental tasks, schema-based and inferential problem solving, in each of the familiar and unfamiliar application domains. We audiotaped and later transcribed all responses.

4 Data, Analysis, and Findings

We first present an example of search for information cues using verbal protocol data that we employed in this research. We then present the way in which we analyzed our verbal protocol data and our findings for schema-based and inferential problem-solving tasks.

4.1 Search for Information Cues Using Verbal Protocol Data

Figure 1 shows the process participant L-4 engaged in while solving Task 2b in the familiar domain. The numbers in the callout boxes indicate the sequence in which the participant made utterances. Based on the utterance in the callout box labeled #1, L-4 first examined the entity type SALES AREA. L-4 then referred to the entity type SALES PERSON (shown in callout box #2).

4.2 Data Analysis

We investigated focus in problem-solving processes using problem-solving transitions. Transitions are movements between different units of problem-solving information identified in the verbal data. We analyzed our data in two ways. In idiographic analyses, using what we term transition graphs (see also Srinivasan & Te'eni, 1995), we examined individual participants' problem-solving processes visually based on movement within the graphical conceptual schema. In subsequent nomothetic analyses, we examined aggregations of the transition data. In line with a number of authors (see, e.g., Kim, Hahn, & Hahn, 2000; Srinivasan & Te'eni, 1995; Zevon & Tellegen, 1982), we then triangulated findings based on our idiographic and nomothetic data to present problem-solving processes. We present our idiographic and nomothetic analyses in turn.

4.2.1 Transition Graphs

The two-dimensional transition graph is a temporal representation of a participant's utterances that reflects their problem-solving process. Each transition on a transition graph (see Figures 2 and 3) represents a building block or a unit of problem solving that involves: 1) understanding the problem, 2) traversing the schema, 3) using the data dictionary to better understand the schema and the problem itself, or 4) assimilating information not provided in the experimental material with that in the schema and data dictionary to solve the problem. When problem solving focused repeatedly on the same material, we included repetitions in our analyses because they reflect uncertainty in the problem-solving process.

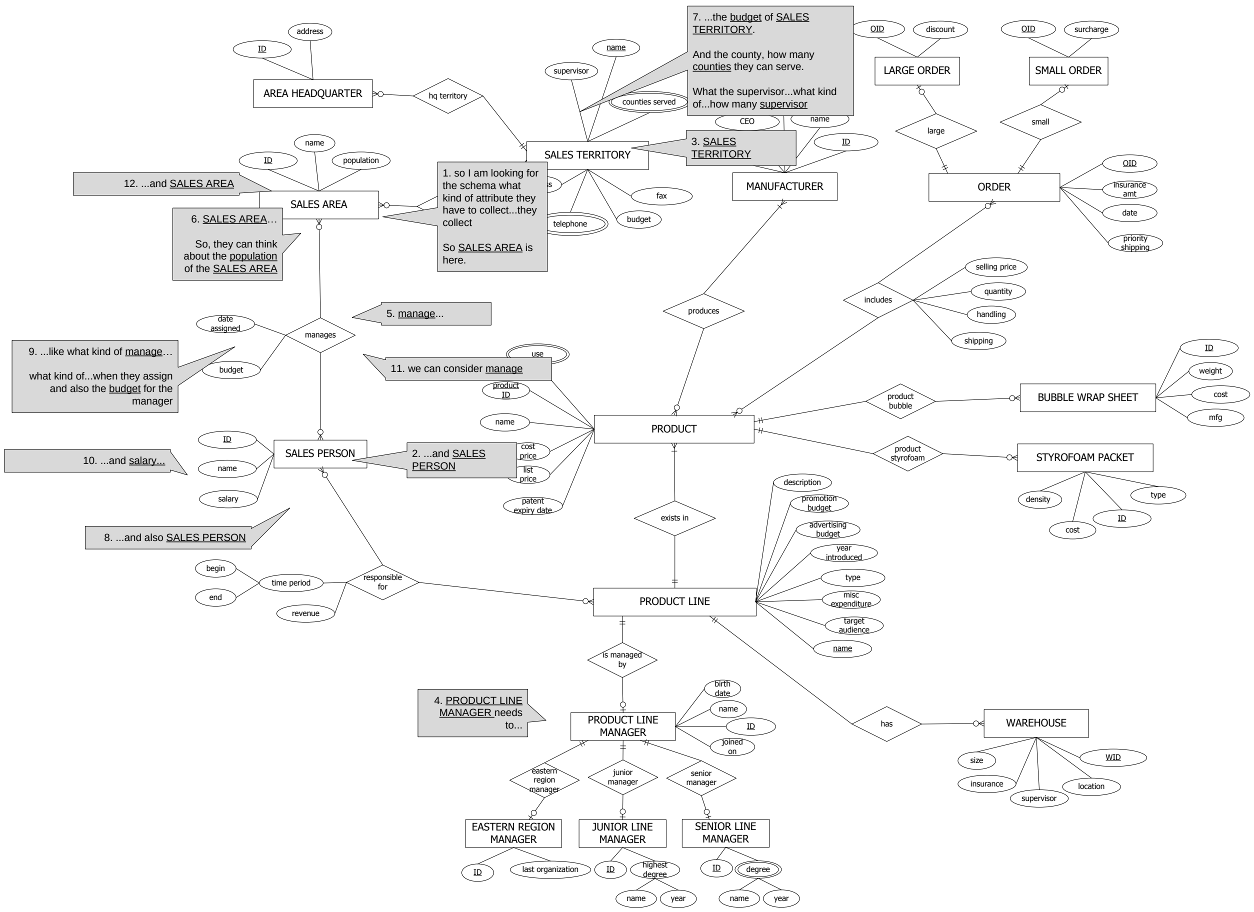


Figure 1. An Example of Schema Traversal for Task 2b in Sales by L-4

Each indicator in the graphs represents a participant's verbalization of the problem statement (✸), reference to the data dictionary (▲), an E/R (○) or an attribute (●), or an utterance outside the schema or the data dictionary (⬠). The numbers in the attribute symbol refer to the attribute referenced by the participant; for example, in Figure 2, graph c, the participant referred to an attribute of the entity type, PRODUCT LINE, denoted by "1", twice, as shown by the appearance of a black circle containing the figure "1" on two occasions. The number in ⬠ (e.g., Figure 3, graph c) refers to an utterance that did not refer to the schema or the data dictionary; it corresponds to the fourth column (i.e., "references other than the schema and data dictionary") of the descriptive data for inferential problem-solving tasks (see Tables E3 and E4). The vertical lines that link objects in the graph reflect the number of transitions needed to address a problem.

4.2.2 Analysis of Transitions

As we show above, we investigate problem solvers' transitions in both ideographic and nomothetic analyses to characterize the search process. We first show the four building blocks of the problem-solving process in the idiographic analyses based on selected transition graphs from each treatment group. We then aggregate the findings from the idiographic analysis to present a nomothetic analysis of the transitions of each participant group.

In our idiographic analyses, we chose a representative participant in each of our treatment groups based on analyzing the median number of transitions. Representative participants are those whose problem-solving processes most closely reflected the median number of transitions in each of the familiar and unfamiliar application domains. For example, the median number of transitions for L-ISDK participants in completing Task 1a was three and four for the familiar and unfamiliar application domains, respectively. Here, we chose L-3 as the representative participant because L-3 had the same number of transitions in the solution. We examined the problem-solving processes of high and low ISDK participants on four problem-solving tasks (two schema-based and two inferential) in both familiar and unfamiliar application domains. Table 5 presents the problem-solving solutions we examined.

Table 5. Transition Graphs Analyzed

Task category	Task	Subjects	Figure
Schema-based problem solving	1a	H-6 and L-3	Figure 2
	1b	H-5 and L-2	Figure 3
Inferential problem-solving	2a	H-3 and L-5	Figure 4
	2b	H-5 and L-3	Figure 5

We present the ideographic analyses of the selected respondents in Sections 4.3.1 and 4.4.1 for schema-based and inferential problem-solving tasks, respectively.

In our nomothetic analyses, we present our analysis of aggregated transition data. Appendix E presents the utterances that the respondents made in schema-based and inferential problem-solving tasks. We next presented these utterances as cues in the schema (entity type, attribute, or relationship labels) or the data dictionary. We provide the intermediate findings in Appendix F, which presents the raw transition data that comprises the number of problem-solving transitions in which our H- and L-ISDK participants engaged in each of the familiar and unfamiliar application domains. We then aggregated that data to conduct the analyses we present in Sections 4.3.2 and 4.4.2 for schema-based and inferential problem-solving tasks, respectively.

4.3 Findings for Schema-based Problem-solving Tasks

Tables E1 and E2 present descriptive data for each of the participants in completing the two schema-based problem-solving tasks in the familiar and unfamiliar domains, while Table F1 summarizes the transitions that each participant made. In this section, we present the findings for the idiographic analyses and those for the aggregate transition analyses.

4.3.1 Idiographic Analyses for Schema-based Problem-solving Tasks

Figures 2 and 3 present transition graphs for schema-based Tasks 1a and 1b, respectively. The region outlined in grey shows the most parsimonious problem-solving process for each of the tasks. Task 1a required referencing a single entity type, PRODUCT LINE, in the familiar application domain and BORE HOLE SITE in the unfamiliar application domain. Figure 2 shows the transition graphs for H-6 and L-3 in each of the application domains on Task 1a. The four graphs show that only the graph of H-6 in the familiar application domain (top left) shows truly focused (i.e., parsimonious) problem solving. From the viewpoint of ISDK, in the familiar domain, H-6 made one transition, while L-3 displayed a further two transitions by referencing an unnecessary entity type, PRODUCT. This observation readily shows, therefore, that L-3 engaged in less-focused problem solving than H-6.

In the unfamiliar application domain, H-6 and L-3 made three and four transitions, respectively, with both participants making an unnecessary reference to the entity type BORE HOLE. The processes are essentially similar except that L-3 made a further transition to reference the problem statement on a second occasion and, thus, exhibited uncertainty. Thus, based on analyzing our representative participants in completing Task 1a, we found that problem solvers with high ISDK used more focused problem-solving processes than problem solvers with low ISDK on the schema-based problem-solving task.

From the viewpoint of ADK, H-6's problem-solving process in the unfamiliar application domain was less focused than in the familiar application domain: the participant made an unnecessary transition to a second entity type, BORE HOLE (Figure 2, graphs a and b). Further, L-3 experienced somewhat greater uncertainty in the unfamiliar application domain by making a further transition to reread the problem statement.

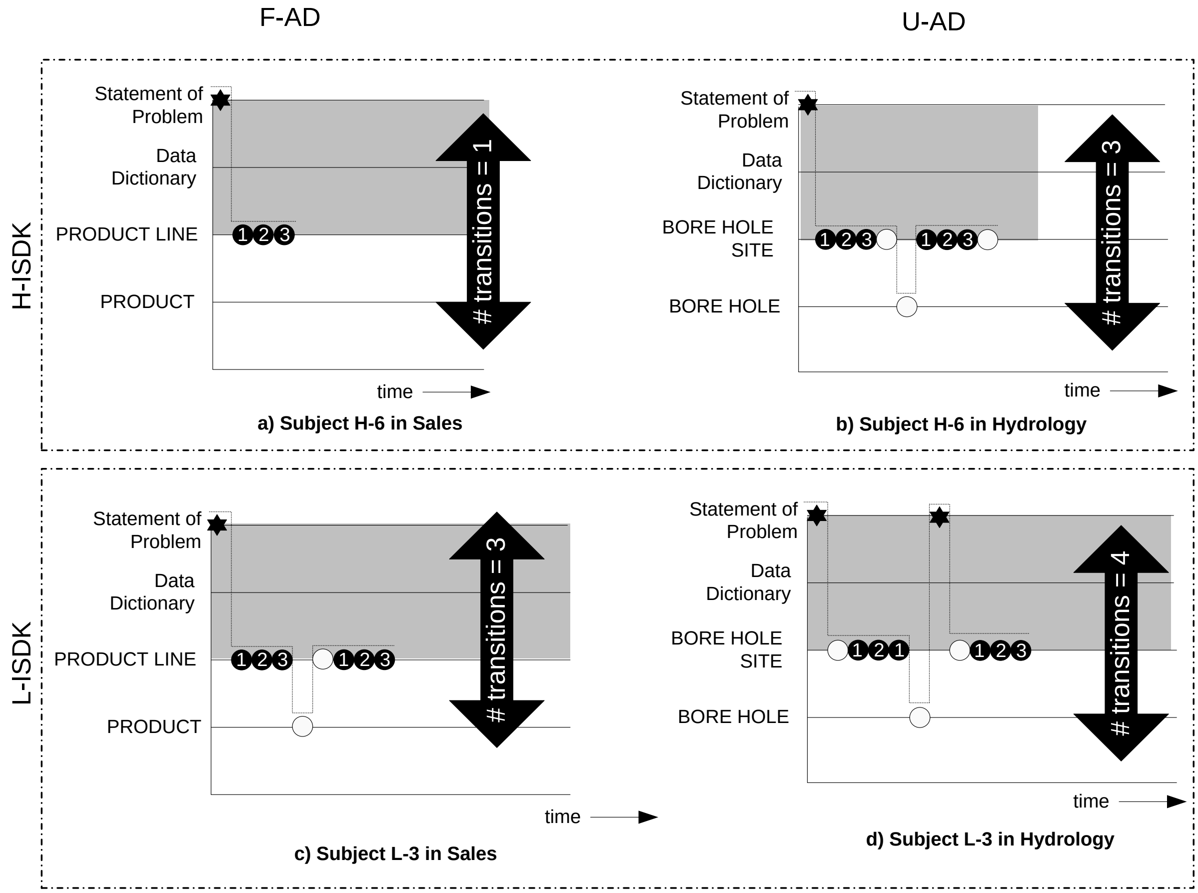


Figure 2. Transition Graphs of Participants for the Schema-based Problem-solving Task 1a

The majority of processes that our participants exhibited reflect findings that show that the participants had greater difficulty solving schema-based problem-solving tasks in unfamiliar domains: four out of six H-ISDK participants and five out of six L-ISDK participants followed a similar pattern (see Table F1). Thus, based on analyzing the data from our participants in completing Task 1a, we found that problem solvers used more-focused problem-solving processes in the familiar than in the unfamiliar application domain on the schema-based problem-solving task.

We now turn to the transition graphs for Task 1b (see Figure 3). Addressing Task 1b requires referencing two entity types: PRODUCT and MANUFACTURER in the familiar application domain and LITHOLOGY and BORE HOLE in the unfamiliar application domain. Figure 3 shows the transition graphs for H-5 and L-2 in each of the application domains. From the viewpoint of ISDK, the four transition graphs show that H-5 made four transitions and referenced an entity type outside the parsimonious solution. On the other hand, while L-2's solution was within the parsimonious region from the viewpoint of referencing relevant entity types, L-2 made several transitions between entity types in this region and, thus, displayed considerable uncertainty (total of six transitions). In the unfamiliar application domain, H-5 referenced three entity types/relationships (two of which were unnecessary), failed to reference the necessary entity type (LITHOLOGY), and made seven transitions. L-2 referenced four entity types/relationships (three of which were unnecessary), also failed to reference the essential LITHOLOGY, and made nine transitions. Overall, these observations show that L-2 engaged in less-focused problem solving than H-5. Thus, we can see again that problem solvers with high ISDK used more-focused problem-solving processes than problem solvers with low ISDK.

From the viewpoint of ADK, H-5 referenced two entity types/relationships outside the parsimonious region in the unfamiliar application domain, failed to reference LITHOLOGY, and referenced the data dictionary on two occasions, which resulted in a total of seven transitions compared with four in the familiar application domain. Similarly, L-2 made transitions to four entity types/relationships in the unfamiliar application domain and failed to reference LITHOLOGY. L-2 also referenced several attributes. L-2 made nine transitions in the unfamiliar application domain compared with six in the familiar application domain.

These findings illustrating the fact that participants engage in less focused problem-solving for schema-based problem-solving tasks in unfamiliar domains are reflected in the processes exhibited by the majority of our participants (see Table E1). All six H-ISDK participants and five out of six L-ISDK participants engaged in more-focused problem-solving processes in the familiar application domain. We can see, therefore, that problem solvers used more-focused problem-solving processes in the familiar than in the unfamiliar application domain.

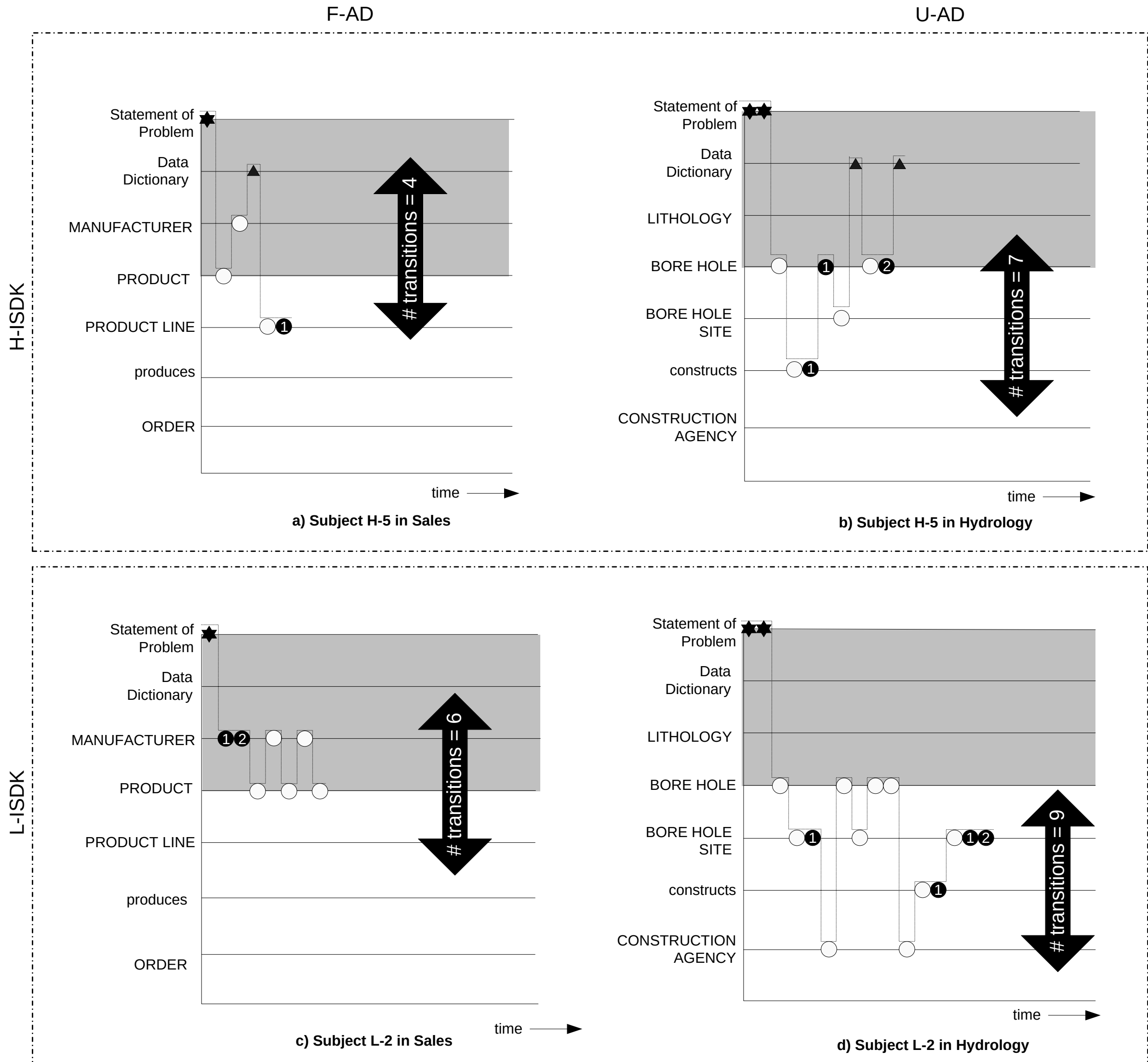


Figure 3. Transition Graphs of Representative Participants for the Schema-based Problem-solving Task 1b

4.3.2 Nomothetic Analyses for Schema-based Problem-solving Tasks

We first evaluated our data on schema-based problem-solving tasks for the effect of ISDK and then of ADK. Table 6 presents the findings of our transition analysis of the effects of ISDK on the problem-solving process on Tasks 1a and 1b for both the familiar (F-AD) and the unfamiliar (U-AD) application domains. In both application domains, H-ISDK participants consistently engaged in more-focused problem solving than L-ISDK participants. Hence, our findings suggest that problem solvers with high ISDK use more-focused problem-solving processes than problem solvers with low ISDK.

Table 7 presents the findings of our transition analyses of the effect of ADK on the problem-solving processes on Tasks 1a and 1b. Both H- and L-ISDK participants consistently engaged in more-focused problem solving in the familiar than in the unfamiliar application domain.

Table 6. Transition Analyses of Effects of IS Domain Knowledge on Problem-solving Focus on Schema-based Problem-solving Tasks 1a and 1b

ADK	ISDK	Median number of transitions	Findings
Task 1a			
F-AD	H-ISDK	2	H-ISDK participants were more focused than L-ISDK participants
	L-ISDK	3	
U-AD	H-ISDK	3	H-ISDK participants were more focused than L-ISDK participants
	L-ISDK	4	
Task 1b			
F-AD	H-ISDK	4	H-ISDK participants were more focused than L-ISDK participants
	L-ISDK	7	
U-AD	H-ISDK	6.5	H-ISDK participants were more focused than L-ISDK participants
	L-ISDK	9	

Table 7. Transition Analyses of Effects of Application Domain Knowledge on Problem-solving Focus on Schema-based Problem-solving Tasks 1a and 1b

ISDK	ADK	Median number of transitions	Findings
Task 1a			
H-ISDK	F-AD	2	Participants were more focused in F-AD than U-AD
	U-AD	3	
L-ISDK	F-AD	3	Participants were more focused in F-AD than U-AD
	U-AD	4	
Task 1b			
H-ISDK	F-AD	4	Participants were more focused in F-AD than U-AD
	U-AD	6.5	
L-ISDK	F-AD	7	Participants were more focused in F-AD than U-AD
	U-AD	9	

4.4 Findings for Inferential Problem-solving Tasks

We now focus on the problem-solving processes that the participants used in completing inferential problem-solving tasks. Tables E3 and E4present the specific transitions that each of the participants made on the two inferential problem-solving tasks in the familiar and unfamiliar application domains, while Table F2 summarizes the transitions that each participant made. Further, Table F3 presents details of references to material not included in the schema that appear on the transition graphs. Here, we present the findings for the idiographic analyses and for the aggregate transition data. We then analyze the aggregate transitions to examine the effects of ISDK and ADK on problem solving.

4.4.1 Idiographic Analyses for Inferential Problem-solving Tasks

Figures 4 and 5 present the transition graphs for inferential Tasks 2a and 2b, respectively. Recall that these types of tasks require one not only to "explore" the schema but also to employ knowledge of the application beyond that represented in the schema and the data dictionary. Several factors evidence the exploratory problem solving required to solve these types of tasks: 1) the number of transitions during schema traversal, which indicates the amount of processing conducted; and 2) references to material that did not appear in the experimental material ("outside the schema"). We examine the focus in problem-solving processes for participants H-3 and L-5 based on the transition graphs for Task 2a in Figure 4.

The four transition graphs in Figure 4 show that H-3 engaged more with the task in the familiar application domain (top left) than in the unfamiliar one (top right) and than L-5 in both domains. Further, the transition graphs in the familiar application domain showed more active engagement than those in the unfamiliar domain. From the viewpoint of ISDK, in the familiar application domain, H-3 made 17 transitions, while L-5 made 12 transitions. H-3 referenced three entity types/relationships in the schema and made references outside the schema a further seven times, while L-5 referenced three entity types/relationships and made a further four references outside the schema. Therefore, and as expected, the process H-3 used was considerably more exploratory in nature than the one L-5 used.

A similar situation holds for H-3 and L-5 in the unfamiliar application domain. H-3 made eight transitions, while L-5 made seven; H-3 made references to four entity types/relationships and two references outside the schema, and L-5 made references to two entity types and zero references outside the schema. Overall, these observations show that H-3 engaged in more-exploratory problem solving than L-5. Thus, based on analyzing our representative participants in completing Task 2a, we found that problem solvers with high ISDK used more exploratory problem-solving processes than problem solvers with low ISDK on the inferential problem-solving task.

One can see even more marked differences in problem-solving processes when one considers the role of ADK. H-3 referenced three entity types/relationships and made seven references outside the schema in the familiar application domain compared with four entity types/relationships and two outside references in the unfamiliar domain. L-5 referenced three entity types/relationships and made four references outside the schema in the familiar domain and two entity types and zero outside references in the unfamiliar application domain.

These findings illustrating the fact that participants explore more in the familiar than in the unfamiliar application domain, are reflected in the processes exhibited by the majority of our participants (see Table F2). Five out of six H-ISDK participants and all six L-ISDK participants engaged in more-exploratory problem-solving processes in the familiar application domain compared with the unfamiliar.

We now turn to the transition graphs for Task 2b (see Figure 5). From the viewpoint of ISDK, in the familiar application domain, H-5 made 16 transitions, while L-3 made 10. H-5 referenced five entity types/relationships in the schema and made one reference outside the schema, while L-3 referenced three entity types/relationships but made no references outside the schema. In the unfamiliar application domain, H-5 and L-3 made ten and eight transitions, respectively. H-5 referenced three entity types/relationships and made one outside reference. Further, L-3 referenced three entity types/relationships and made zero references outside the schema. Therefore, we can see again that problem solvers with high ISDK used more exploratory problem-solving processes than problem solvers with low ISDK.

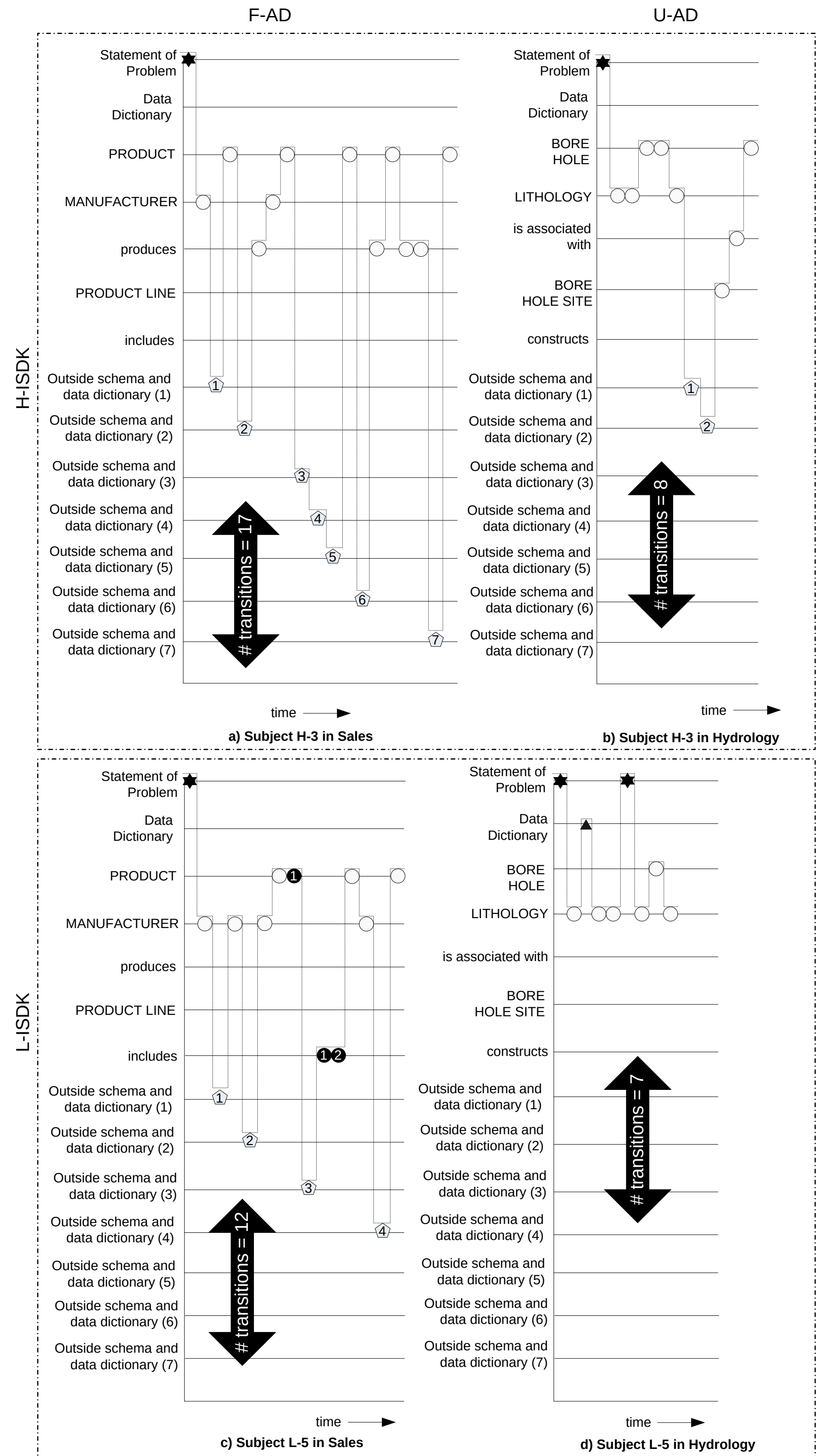


Figure 4. Transition Graphs of Representative Participants on Inferential Problem-solving Task 2a

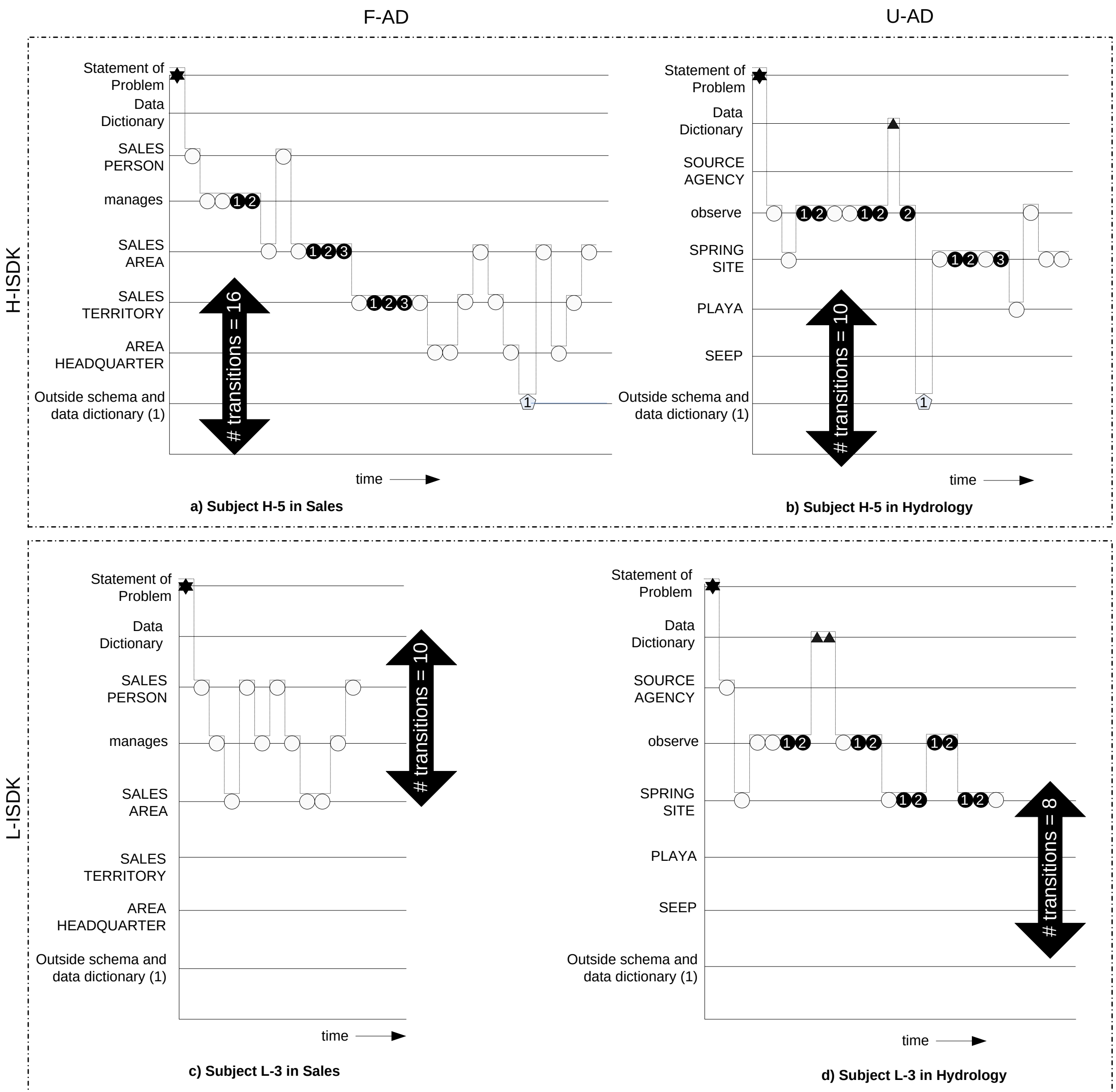


Figure 5. Transition Graphs of Representative Participants on Inferential Problem-solving Task 2b

From the viewpoint of the role of ADK, H-5 made more transitions in the familiar than in the unfamiliar application domain (16 compared to 10), referred to more entity types/relationships (5 compared to 3), and made one outside reference in each application domain. Therefore, H-5's process was more exploratory in the familiar than in the unfamiliar application domain. L-3 made more transitions in the familiar than in the unfamiliar application domain (10 compared to 8), referred to the same number of entity types/relationships (3), and made the same number of outside references (0). As such, based on transition data analysis, L-3's search process was somewhat more exploratory in the familiar than the unfamiliar application domain.

These findings, which suggest that participants explore more in familiar than in the unfamiliar domains, are reflected in the processes exhibited by the majority of our participants (see Table F2). All six H-ISDK participants and five out of six L-ISDK participants engaged in more exploratory problem-solving processes in the familiar than in the unfamiliar application domain.

4.4.2 Nomothetic Analyses for Inferential Problem-solving Tasks

We present our nomothetic evaluation of the effect of ISDK and ADK on inferential problem-solving processes. Table 8 presents the findings of our transition analysis of ISDK's effects on the problem-solving process for our two inferential problem-solving tasks, Tasks 2a and 2b, for both the familiar and unfamiliar application domains. In both application domains, H-ISDK participants consistently engaged in more exploratory problem solving than L-ISDK participants. Hence, our findings suggest that, for inferential problem-solving tasks, problem solvers with high ISDK use more exploratory problem-solving processes than problem solvers with low ISDK.

Table 8. Transition Analyses of Effects of IS Domain Knowledge on Problem-solving Focus on Inferential Problem-solving Tasks 2a and 2b

ADK	ISDK	Median number of transitions	Findings
Task 2a			
F-AD	H-ISDK	16	H-ISDK participants explored more than L-ISDK participants
	L-ISDK	12.5	
U-AD	H-ISDK	9.5	H-ISDK participants explored more than L-ISDK participants
	L-ISDK	7	
Task 2b			
F-AD	H-ISDK	14	H-ISDK participants explored more than L-ISDK participants
	L-ISDK	11	
U-AD	H-ISDK	9	H-ISDK participants explored more than L-ISDK participants
	L-ISDK	8	

Table 9 presents the findings of our transition analysis of ADK's effects on the problem-solving process for Tasks 2a and 2b. Both H- and L-ISDK participants consistently engaged in more exploratory problem solving in the familiar than in the unfamiliar application domain. Hence, our findings suggest that, for inferential problem-solving tasks, problem solvers use more exploratory problem-solving processes in familiar than in unfamiliar application domains.

Table 9. Transition Analyses of Effects of Application Domain Knowledge on Problem-solving Focus on Inferential Problem-solving Tasks 2a and 2b

ISDK	ADK	Median number of transitions	Findings
Task 2a			
H-ISDK	F-AD	16	Participants explore more in F-AD than U-AD
	U-AD	9.5	
L-ISDK	F-AD	12.5	Participants explore more in F-AD than U-AD
	U-AD	7	
Task 2b			
H-ISDK	F-AD	14	Participants explore more in F-AD than U-AD
	U-AD	9	
L-ISDK	F-AD	11	Participants explore more in F-AD than U-AD
	U-AD	8	

5 Theoretical Discussion and Implications

Our research addresses a widely acknowledged though not widely studied issue in IS: that of the role of the application domain in IS problem solving. We address this issue by investigating how both IS and application domain knowledge (ISDK and ADK, respectively) contribute to the process of understanding conceptual schemas. Our research question was: how do IS and application domain knowledge influence the way in which individuals solve conceptual schema problem-solving tasks? Specifically, we examined how IS and application domain knowledge support deep-level understanding of both schema-based and inferential problem-solving tasks in familiar and unfamiliar application domains by examining the processes individuals use to identify information cues.

In this section, we explain theoretically our findings based on the cognitive psychology literature on problem solving. We then discuss our study's contributions to the literature and its implications for future research, for teaching, and for practice.

5.1 Theoretical Explanation of Findings

While researchers have conducted some research into the role of the application domain in addressing conceptual schema problem-solving tasks, it has been largely atheoretical. One may explain our research using the paradigm of humans as information processing systems (HIPS), which views information (the stimulus) as entering the mind and being processed in a series of ordered stages in short-term memory with further information being retrieved from long-term memory as appropriate and with the results of cognitive processing being output or stored in long-term memory (Newell & Simon, 1972). The primary foci of this literature are the cognitive processes and/or strategies that problem solvers use and the role of ISDK and ADK in those processes. Below, we examine the roles of ISDK and ADK in solving schema-based problem-solving tasks followed by inferential problem-solving tasks.

5.1.1 Schema-based Problem-solving Tasks

We characterize schema-based problem-solving tasks as structurable because one can reformulate them using a structuring methodology (Smith, 1988) into tasks that have a well-defined initial state, a clearly-defined goal state, and a well-defined, constrained set of transformation functions to guide a single well-defined optimal solution process (Greeno 1978; Sinnott, 1989; Voss & Post, 1988). Although one may solve schema-based problem-solving tasks using knowledge represented in the schema alone, that information needs to be transformed. Hence, one can characterize these types of tasks as being structurable in nature. We first examine the role of ISDK followed by that of ADK.

With regard to the role of ISDK in solving schema-based problem-solving tasks, prior research that has examined the role of specialized knowledge suggests that problem solvers with high levels of domain knowledge tend to use their knowledge of the problem domain (in this case, ISDK) to guide their search for data to understand the structure of the task at hand (Shanteau, 1992). Problem solvers with a low level of domain knowledge, on the other hand, are heavily influenced by the surface features of the task. Further, they tend to focus on the information that is most readily available (Biggs, Mock, & Watkins, 1988; Bouwman, 1984) and, therefore, are likely to engage in problem solving that is less focused than that of problem solvers with greater domain knowledge (Chi et al., 1981).

Because the fact that all the required information is available in the schema guides one's efforts in solving schema-based problem-solving tasks, we expect that problem solvers will, in general, engage in focused problem solving. Participants with low ISDK may, however, be less certain about the information they need to solve such a task. Hence, we expect participants with high ISDK to engage in processes that are more focused than those with low ISDK. Researchers often refer to more-focused search as "depth-first" based on the fact that the problem solver concentrates on the issue at hand rather exploring the problem space as a whole as in breadth-first search (Chi & Glaser, 1982). Therefore, we propose:

> **Proposition 1:** When solving structurable schema-based problem-solving tasks, problem solvers with high ISDK use more-focused problem-solving processes than problem solvers with low ISDK.

With regard to the role of ADK in solving schema-based problem-solving tasks, because the information for solving a schema-based problem-solving task is not available directly in the schema, other application-related information might aid in task solution. Pirolli, Card, and Van Der Wege (2001), who found that local cues (e.g., text labels presenting schema semantics) influence visual attention (e.g., the lower the strength of local cues, the greater the exploratory search) support this argument.

We investigated the role of additional knowledge in the form of ADK. Our theoretical analyses suggest that, for a structurable task, ADK may play a role in addition to that played by ISDK. That is, ADK may also play a role in solving tasks, a notion that prior research supports (Khatri et al., 2006). Hence, participants who are less familiar with the application domain engage in less-focused search than those who are more familiar with the application domain. Therefore, we propose:

Proposition 2: When solving structurable schema-based problem-solving tasks, problem solvers use more-focused problem-solving processes in the familiar than in the unfamiliar application domains.

5.1.2 Inferential Problem-Solving Tasks

We characterize inferential problem-solving tasks as ill-structured because the situation they address is not directly represented in the conceptual schema, the initial and goal states are vaguely defined or unclear (Voss & Post, 1988), and there are multiple solutions and solution paths or no solution at all (Kitchner, 1983). In further contrast to schema-based problem-solving tasks, the schema alone does not guide the solution process; rather, it is unconstrained and, therefore, unclear. Solving inferential problem-solving tasks may, therefore, result in multiple solutions and/or solution processes due to the lack of guidance in the solution process. In effect, although problem solvers draw on their knowledge of both ISDK and ADK, an inferential problem-solving task foregrounds ADK rather than ISDK. Bodart et al. (2001), who view inferential problem-solving tasks as serving to evaluate participants' "elaborative and inferential reconstruction effects associated with deep processing of [application] domain semantics" (pp. 399-400), supports this view of the importance of the application domain in solving inferential problem-solving tasks. We first examine the role of ISDK followed by that of ADK.

With regard to the role of ISDK in solving inferential problem-solving tasks, prior research suggests that, with ill-structured tasks, problem solvers with greater domain knowledge tend to engage in breadth-first problem-solving strategies to ensure they do not close constraints on their problem solving before establishing that they can reach a viable solution (see, e.g., Adelson & Soloway, 1985; Greeno, 1978; Jeffries et al., 1981; Rist, 1989). Further, to ensure that they do not overlook salient information, problem solvers with greater knowledge also set their information filters lower and, thereby, accept more noise or irrelevant information (Pirolli & Card, 2005). Therefore, problem solvers with greater ISDK explore the problem space more extensively than those with lower ISDK. Thus, we expect that participants with high ISDK manifest processes that are more exploratory than those with low ISDK. Therefore, we propose:

Proposition 3: When solving ill-structured inferential problem-solving tasks, problem solvers with high ISDK use more exploratory problem-solving processes than problem solvers with low ISDK.

With regard to the role of ADK in solving inferential problem-solving tasks, our argument that knowledge of the application domain results in participants setting their information filters lower suggests that participants who are more familiar with the application domain engage in more exploratory search (see Pirolli & Card, 2005). Further, because problem solvers set the information filters lower in such circumstances, they can better encode new information when familiar with the application domain (see Egan & Schwartz, 1979; Jeffries et al., 1981; Voss, Vesonder, & Spilich, 1980). In turn, this situation requires problem solvers to engage in greater exploratory search (Johnson & Russo, 1984). Hence, we expect that participants engage in more exploratory search in familiar than in unfamiliar application domains. Therefore, we propose:

Proposition 4: When solving ill-structured inferential problem-solving tasks, problem solvers use more exploratory problem-solving processes in the familiar than in the unfamiliar application domains.

Hence, we postulate that both IS and application domain knowledge influence problem-solving processes when solving ill-structured inferential problem-solving tasks. That is, participants use exploratory (breadth-first) processes when they have high IS domain knowledge or are familiar with the application domain.

5.1.3 Contributions

Table 10, which summarizes the theoretical explanation of our findings, shows how the nature of the problem-solving process depends on the extent of structure in the conceptual schema problem-solving task (i.e., whether the task is structurable or unstructured in nature) and the levels of ISDK and ADK.

Our research makes several contributions to the literature by providing insights into how ISDK and ADK affect conceptual schema problem solving. First, we provide theoretical foundations for explaining the role of ADK in IS problem solving by examining the cognitive psychology literature on problem solving. We then characterize schema-based and inferential problem-solving tasks as structurable and ill-structured tasks, respectively, based on the extent of structure in the problem-solving task.

Table 10. Summary of Effects of IS and Application Domain Knowledge on Different Types of Conceptual Schema Problem-solving Tasks

IS task	Type of IS domain knowledge needed for problem solving	Type of application knowledge needed for problem solving	
		High application domain knowledge	Low application domain knowledge
Schema-based problem-solving tasks • All information for problem solving available in the conceptual schema	High IS domain knowledge	• IS domain knowledge guides focused search. • Fewer transitions to arrive at the goal state due to application domain knowledge.	• IS domain knowledge guides focused search. • More transitions to arrive at the goal state because of lack of application domain knowledge.
	Low IS domain knowledge	• Lack of IS domain knowledge results in less-focused, more-exploratory search. • Fewer transitions to arrive at the goal state due to application domain knowledge.	• Lack of IS domain knowledge results in less-focused, more-exploratory search. • More transitions to arrive at the goal state due to lack of application domain knowledge.
Inferential problem-solving tasks • Use of information beyond that provided in the conceptual schema	High IS domain knowledge	• IS domain knowledge guides exploratory search. • Higher application domain knowledge results in lower filter for irrelevant information.	• IS domain knowledge guides exploratory search. • Lower application domain knowledge results in higher filter for irrelevant information.
	Low IS domain knowledge	• Lack of IS domain knowledge results in focused search. • Higher application domain knowledge results in lower filter for irrelevant information.	• Lack of IS domain knowledge results in focused search. • Lower application domain knowledge results in higher filter for irrelevant information.

Second, we characterize the types of processes participants use to solve conceptual schema problem-solving tasks as follows: 1) focused (depth-first) processes when IS domain knowledge is high, the application domain is familiar, and the information required for problem solving is available in the conceptual schema; and 2) exploratory (breadth-first) processes when IS domain knowledge is high, the application domain is familiar, and the necessary information does not appear in the schema.

Third, we observe from our transition graphs that, when solving schema-based problem-solving tasks, our participants exhibited limitations in cognitive capacity (i.e., stress; see, e.g., Tarafdar, Tu, & Ragunathan, 2010). Such participants displayed less-focused problem-solving processes that one could regard as somewhat haphazard. The existence of stress, which is particularly salient in our visual transition graphs, was exacerbated when IS domain knowledge was low and the application domain was unfamiliar. The transition graph of the low IS domain knowledge participant on Task 1b in the unfamiliar application domain illustrates this situation well (see Figure 3). Thus, in identifying problem-solving stress, we further exemplify the insights our research provides into how lack of ISDK and ADK affect problem solving.

Our study has the following limitations. First, we conducted our investigation using students who were relatively inexperienced in using real-world conceptual schemas. We characterize them as new recruits entering the practical role of data analyst. Note, however, that the difference between students and professionals is not always clear. For example, a study on maintaining UML diagrams found no differences in performance of undergraduate/graduate students and junior/intermediate professional consultants (Arisholm & Sjøberg, 2004).

Second, while researchers have questioned verbal protocol data on several issues (see, e.g., Nisbett & Wilson, 1977, for a detailed analysis), it remains the accepted way of collecting process data. It is, for example, a far better approach than using retrospective reports or various types of self-reported data.

Third, due to the density of verbal protocol data, researchers typically use a small number of participants (typically "between 2 and 20") in their experiments (Todd & Benbasat, 1987, p. 501). Appendix G presents the sample size of several such IS research studies. Our study, with 12 participants, is in the mid-range. Further, our sample is large enough to obtain meaningful results. Further again, we strengthen our analysis by presenting both idiographic and nomothetic analyses of the data (see also Barley, 1990; DeSanctis & Poole, 1997; Hungerford et al., 2004; Kim et al., 2000). We also show that our findings are consistent across multiple tasks.

5.2 Implications of the Findings

In this section, we present the implications of our findings for research, teaching, and practice.

5.2.1 Implications for Research

Our findings have several implications for researchers. From the viewpoint of the role of the application domain in IS problem solving, we present theory that explains the role of the IS and application domains in solving a range of IS problems, which should prove valuable to the growing numbers of researchers who are interested in pursuing the role of the application domain in IS. First, research needs to be conducted to establish the boundaries to the propositions presented. Our findings are generalizable only to novice IS conceptual modelers, and future research should be undertaken to examine the applicability of our findings to the general population of IS professionals who have years of experience both in the IS domain and in multiple application domains.

Second, we need further research to determine whether the effects of task structure observed here hold for similar types of problems. We also need research to characterize how task structure affects problem-solving processes in general and process diagram comprehension in particular (e.g., by examining problem solving using data-flow diagrams).

Third, our research focuses on characterizing problem-solving processes only for ERA in ER and EER models. We need further research into using other conceptual models such as the class diagrams of UML. We note, again, however, that the ER model is by far the most popular data model in practice (Davies et al., 2004). Fourth, we need further research into information search for conceptual schema-understanding tasks that result from semantic ambiguities of ontologically unsound representations (see, e.g., Burton-Jones & Weber 1999; Shanks et al., 2008). Researchers should conduct such research in the context of application domain knowledge.

Fifth, while this study extends our knowledge of schema understanding, we need further protocol analysis studies in the context of schema development. Srinivasan and Te'eni (1995), for example, found that specific strategies for building a conceptual schema affect the quality of the resulting representation. We need to extend their research to include the roles of both ISDK and ADK.

Sixth, future research needs to explore how to represent application domain knowledge in the schema. For example, prior research in process modeling (Mendling, Reijers, & Recker, 2010) suggests that the specific activity constructs used have significant impact on the perceived ambiguity and usefulness of the labels. Future research needs to explore an approach for labeling the constructs in conceptual schema and, thus, embed application domain knowledge in the schema. Future research could, for example, examine how incorporating graphical icons in additional textual annotations influences search for schema understanding (Mayer, 1989; Paivio, 1991).

5.2.2 Implications for Teaching

Our findings on conceptual schema problem solving have several implications for instruction in conceptual modeling. We found that participants with high ISDK in the familiar application domain undertook depth-first problem solving for schema-based problem-solving tasks and breadth-first problem solving for inferential problem-solving tasks. Hence, instruction needs to focus on the application domain and extent of structure in the tasks that students/trainees undertake as they try to better understand conceptual schemas.

Second, our research indicates that introductory courses in conceptual modeling should focus initially on familiar application domains; that is, domains in which the students are likely to be knowledgeable. In this way, one can guide students toward using an effective process before addressing tasks that are less structured.

Third, examining students' interaction with conceptual schemas could provide useful insights into their developing expertise. For example, problem-solving processes could provide subtle feedback to instructors regarding the learning that students experience. Further, instructors could evaluate proficiency in problem solving by analyzing problem-solving processes.

5.2.3 Implications for Practice

From a practical perspective, our research has several implications for managing and training IT personnel and for developing tools to support conceptual schema problem-solving processes.

First, for schema-based problem-solving tasks, our research findings suggest that both IS personnel and users will perform better when they have both types of domain knowledge. Note, however, that prior research suggests that, for ill-structured problems, ADK can create biases and lead knowledgeable subjects to search narrowly, which decreases their chances of finding an appropriate solution (Berry, 1995; Buchman & Ekadharmawan, 2009; Hadar et al., 2014; Wiley, 1998). Recognizing that the scope of our research is the well-structured problem area of conceptual schema understanding, management should seek to ensure that they assign both IT and business personnel to applications that match not only their ISDK but also their ADK. That is, personnel knowledgeable in a given application domain are best assigned to projects focusing on that domain. For example, IT personnel with ADK should be able to employ appropriate search processes and, thereby, better align business and IT objectives. In the event that one assigns personnel with little knowledge of the application domain to a project, one can expect the project to result in increased use of resources and be less effective than might otherwise be the case.

Second, in addition to organizations providing systematic training for IS personnel on ISDK and tool knowledge, they should also consider providing systematic training on ADK. One can gauge the effectiveness of training in ISDK and ADK by examining the processes that IS personnel employ.

Third, the growing body of evidence pointing to the importance of application domain knowledge in certain types of IS problem solving suggests that tool builders should investigate ways to incorporate characteristics of the application domain into their tools. Tools could potentially support the schema understanding process through, perhaps, using domain-specific modeling patterns and templates, which would help reduce the time and effort expended. Domain-specific modeling languages (DSMLs), for example, involve two stages: 1) defining conceptual constructs to address a specific application domain and 2) implementing the specification as UML extensions in the form of stereotypes (Lagarde, Espinoza, Terrier, & Gerard, 2007). Incorporating characteristics of the application domain into tools can, therefore, aid in defining conceptual constructs that address specific application domains.

One could also design tools to support different types of problem-solving processes by presenting the schema in a way that is appropriate for the task at hand. For example, to support breadth-first search on inferential problem-solving tasks, one could show IS personnel all entity types (i.e., the entire schema without the attributes). On the other hand, to support depth-first search on schema-based problem-solving tasks, one could show IS personnel a subset of the schema that is relevant for addressing the task.

6 Conclusion

The role of the application domain is an issue that research on IS problem solving has largely neglected. In this study, we explored how IS and application domain knowledge (ISDK and ADK, respectively) each influence the solution of conceptual schema problem-solving tasks. Analyses of problem-solving processes for schema-based and inferential problem-solving tasks revealed that the problem-solving processes that result with better ISDK and/or better ADK differ according to the extent of structure in the task. Problem solvers engage in focused (depth-first) problem solving when solving structurable schema-based problem-solving tasks. However, they engage in exploratory (breadth-first) problem solving when solving ill-structured inferential problem-solving tasks. For the two types of conceptual schema problem-solving tasks examined in this research, we discovered that ADK and ISDK have similar effects on problem-solving processes. Therefore, this research provides insight into how ISDK and ADK affect IS problem solving.

Our study contributes to the growing recognition and examination of the role of ADK in IS problem solving and provides guidance for training students and practitioners by, for example, acknowledging the need for ADK, recommending the use of effective problem-solving processes, and constructing tools to support the conceptual schema understanding process. We hope that focusing attention on how both ISDK and ADK affect problem solving will aid in moving both IS research and practice forward.

Acknowledgments

We are indebted to Senior Editor, Maung Sein, and two anonymous reviewers for their counsel, feedback, and insightful suggestions on the paper as it evolved through the review process. We also like to thank colleagues at Indiana University and the University of Queensland for their constructive comments on earlier versions of the paper.

References

Adelson, B., & Soloway, E. (1985). The role of domain experience in software design. *IEEE Transactions on Software Engineering*, *11*(11), 1351-1360.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*(4), 369-406.

Anderson, J. R. (1996). *The architecture of cognition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Arisholm, E., & Sjøberg, D. I. K. (2004). Evaluating the effect of a delegated versus centralized control style on the maintainability of object-oriented software. *IEEE Transactions on Software Engineering*, *30*(8), 521-534.

Barley, S. R. (1990). Images of imaging: Notes on doing longitudinal field work. *Organization Science, 1*(3), 220-245.

Bera, P., Burton-Jones, A., & Wand, Y. (2009). The effect of domain familiarity on modelling roles: An empirical study. In *Proceedings of the Pacific Asia Conference on Information Systems.*

Berry, D. M. (1995). The importance of ignorance in requirements engineering. *Journal of Systems and Software*, *28*(2), 179-184.

Biggs, S. F., Mock, T. J., & Watkins, P. R. (1988). Auditor's use of analytical review in audit program design. *The Accounting Review, 63*(1), 148-161.

Blum, B. A. (1989). A paradigm for the 1990s validated in the 1980s. In *Proceedings of the AIAA Conference* (pp. 502-511).

Bodart, F., Patel, A., Sim, M., & Weber, R. (2001). Should optional properties be used in conceptual modelling? A theory and three empirical tests. *Information Systems Research*, *12*(4), 384-405.

Bouwman, M. (1984). Expert vs. novice decision making in accounting: A summary. *Accounting, Organizations and Society*, *9*(3-4), 1984, 325-327.

Browne, G. J., & Parsons, J. (2012). More enduring questions in cognitive IS research. *Journal of the Association For Information Systems*, *13*(12), 1000-1011.

Buchman, J., & Ekadharmawan, C. (2009). Barriers to sharing domain knowledge in software development practice in SMEs. In *Proceedings of the 3rd International Workshop on Knowledge Collaboration in Software Development* (pp. 2-16).

Burton-Jones, A., & Meso, P. N. (2006). Conceptualizing systems for understanding: An empirical test of decomposition principles in object-oriented analysis. *Information Systems Research*, *17*(1), 38-60.

Burton-Jones, A., & Meso, P. N. (2008). The effects of decomposition quality and multiple forms of information on novices' understanding of a domain from a conceptual model. *Journal of the Association for Information Systems*, *9*(12), 748-802.

Burton-Jones, A., & Weber, R. (1999). Understanding relationships with attributes in entity-relationship diagrams. In *Proceedings of the 20th International Conference on Information Systems* (pp. 214-228).

Chen, P. P. (1976). The entity-relationship model—toward a unified view of data. *ACM Transactions of Database Systems*, *1*(1), 9-36.

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

Chi, M., & Glaser, R. (1984). Problem solving ability. In R. J. Sternberg (Ed*.), Human abilities: An information processing approach* (pp. 227–250). San Francisco: Freeman.

Chi, M. T. H., & Glaser, R. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*, (Vol. 1). Hillsdale, NJ: Erlbaum.

Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM*, *13*(6), 377-387.

Davern, M., Shaft, T., & Te'eni, D. (2012). Cognition matters: Enduring questions in cognitive IS research. *Journal of Association of Information Systems, 13*(4), 273-314.

Davies, I., Green, P., Rosemann, M., & Gallo, S. (2004). Conceptual modelling—what and why in current practice. In *Proceedings of the 23rd International Conference on Conceptual Modeling* (pp. 30-42).

Davies, I., Green, P., Rosemann, M., Indulska, M., & Gallo, S. (2006). How do practitioners use conceptual modeling in practice? *Data & Knowledge Engineering*, *58*(3), 358-380.

Delbaere, M., & Ferreira, R. (2007). Addressing the data aspects of compliance with industry models. *IBM Systems Journal*, *46*(2), 319-334.

DeSanctis, G., & Poole, M. S. (1997). Capturing the complexity in advanced technology use: Adaptive structuration theory. *Organization Science*, *5*(2), 121-147.

Edberg, D., Ivanova, P., & Kuechler, W. (2012). Methodology mashups: An exploration of processes used to maintain software. *Journal of Management Information Systems, 28*(4), 271-304.

Egan, D., & Schwartz, B. (1979). Chunking in recall of symbolic drawings. *Memory and Cognition, 7*(2), 149-158.

Elmasri, R., & Navathe, S. B. (2006). *Fundamentals of database systems* (6th ed.). Boston, MA: Addison Wesley.

Fettke, P. (2009). How conceptual modeling is used. *Communications of the Association for Information Systems, 25,* 571-592.

Geiger, J. (2010). Why do we model? *Information Management, 20*(2), 32-33.

Gemino, A. (1999). *Empirical methods for comparing system analysis modelling techniques* (doctoral thesis). Vancouver, BC: University of British Columbia.

Gemino, A., & Wand, Y. (2003). Evaluating modeling techniques based on models of learning. *Communications of the ACM, 46*(10), 79-84.

Gemino, A., & Wand, Y. (2005). Complexity and clarity in conceptual modeling: Comparison of mandatory and optional properties. *Data and Knowledge Engineering*, *55*, 301-326.

Greeno, J. (1978). Natures of problem-solving abilities. In W. Estes (Ed.), *Handbook of learning and cognitive processes* (pp. 239-270). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hadar, I., Soffer, P., & Kenzi, K. (2014). The role of domain knowledge in requirements elicitation via interviews: An exploratory study. *Requirements Engineering Journal, 19*(2), 143-159.

Hoffer, J. A., Prescott, M. B., & McFadden, F. R. (2007). *Modern database management* (8th ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.

Hungerford, B. C., Hevner, A. R., & Collins, R. W. (2004). Reviewing software diagrams: A cognitive study. *IEEE Transactions in Software Engineering, 30*(2), 82-96.

Jeffries, R., Turner, A. A., Poison, P., & Atwood, M. (1981a). The processes involved in designing software. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 255-283). Hillsdale, NJ: Erlbaum.

Johnson, E. J., & Russo, J. E. (1984). Product familiarity and learning new information. *Journal of Consumer Research*, *11*(1), 542-550.

Johnson, P., Duran, A., Hassebrock, F., Miller, J., Prietula, M., Feltovich, P., & Swanson, D. (1981). Expertise and error in diagnostic reasoning. *Cognitive Science*, *5*, 235-283.

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148-152.

Khatri, V., Vessey, I., Ramesh, V., Clay, P., & Park, S.-J. (2006). Understanding conceptual schemas: Exploring the role of application and IS domain knowledge. *Information Systems Research*, *17*(1), 81-99.

Kim, J., Hahn, J., & Hahn, H. (2000). How do we understand a system with (so) many diagrams? Cognitive integration processes in diagrammatic reasoning. *Information Systems Research*, *11*(3), 284-303.

Kim, Y.-G., & March, S. T. (1995). Comparing data modeling formalisms. *Communications of the ACM, 38*(6), 103-115.

Kitchner, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development, 26*, 222-232.

Koskinen, J. (2010). Experiences from software maintenance seminars: Organizing three seminars with 127 groups. In M. Iskander, V. Kapila, M. A. Karim (Eds.), *Technological developments in education and automation* (pp. 339-344). Netherlands: Springer.

Lagarde, F., Espinoza, H., Terrier, F., & Gérard, S. (2007). Improving UML profile design practices by leveraging conceptual domain models. In *Proceedings of the 22nd IEEE/ACM International Conference on Automated Software Engineering* (pp. 445-448).

Lukyanenko, R., Parsons, J., & Wiersma, Y. F. (2015). The IQ of the crowd: Understanding and improving information quality in structured user-generated content. *Information Systems Research, 25*(4), 669-689.

Malhotra, A., Thomas, J. C., Carrol, J. M., & Miller, L. A. (1980). Cognitive processes in design. *International Journal of Man-Machine Studies, 12*, 119-140.

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

Mayer, R. E. (1997). From novice to expert. In M. Helander, T. Landauer, & P. Prabbu (Eds.), *Handbook of human-computer interaction* (pp. 781-795). London, UK: Elsevier Science.

Mendling, J., Reijers, H. A., & Recker, J. (2010). Activity labeling in process modeling: Empirical insights and recommendations. *Information Systems*, *35*(4), 467-482.

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

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM, 19*(3), 113-126.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255-287.

Petre, M. (1995). Why looking isn't always seeing: Readership skills and graphical programming. *Communications of the ACM, 38*(6), 33-44.

Pirolli, P., Card, S., & Van Der Wege, M. (2000). The effect of information scent on searching information visualizations of large tree structures. In *Proceedings of AVI* (pp. 161-172).

Pirolli, P., & Card, S. K. (2005). The sensemaking process and leverage points for analyst technology. In *Proceedings of the International Conference on Intelligence Analysis.*

Power, D. (2011). Data modeling: The lingua franca of a successful MDM effort. *Information Management, 21*(1), 28-29.

Purao, S., Rossi, M., & Bush, A. (2002). Toward an understanding of the use of problem and design spaces during object-oriented system development. *Information and Organization, 12*(4), 249-281.

Rist, R. S. (1989). Schema creation in programming. *Cognitive Science, 13*(3), 389-414.

Robson, D., Bennett, K. B., Cornelius, B., & Munro, M. (1991). Approaches to program comprehension. *The Journal of Systems and Software, 14*, 79-84.

Saghafi, A., & Wand, Y. (2014). Do ontological guidelines improve understandability of conceptual models? A meta-analysis of empirical work. In *Proceedings of the 47th Hawaii International Conference on System Science* (pp. 4609-4618).

Shaft, T., & Vessey, I. (1995). The relevance of application domain knowledge: The case of computer program comprehension. *Information Systems Research, 6*(3), 286-299.

Shaft, T., & Vessey, I. (1998). The relevance of application domain knowledge: Characterizing the computer program comprehension process. *Journal of Management Information Systems, 15*(1), 51-78.

Shaft, T., & Vessey, I. (2006). The role of cognitive fit in the relationship between software comprehension and modification. *MIS Quarterly, 30*(1), 29-55.

Shanks, G., Nuredini, J., Tobin, D., Moody, D., & Weber, R. (2003). Representing things and properties in conceptual modeling: An empirical investigation. In *Proceedings of the European Conference on Information Systems.*

Shanks, G., Nuredini, J., & Weber, R. (2004). Evaluating conceptual modeling practices: Composites, things, properties. In P. F. Green & M. Rosemann (Eds.), *Business systems analysis with ontologies* (pp. 28-55) Hershey, PA: Idea Group Publishing.

Shanks, G., Tansley, E., Nuredini, J., Tobin, D., & Weber, R. (2002). Representing part-whole relationships in conceptual modeling: An empirical evaluation. In *Proceedings of the International Conference on Information Systems* (pp. 89-100).

Shanks, G., Tansley, E., Nuredini, J., Tobin, D., & Weber, R. (2008). Representing part-whole relations in conceptual modeling: An empirical evaluation. *MIS Quarterly*, *32*(3), 553-573.

Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica, 81*, 75-86.

Sheth, A. (1995). Data semantics: What, where and how? In *Proceedings of the 6th IFIP Working Conference on Data Semantics* (pp. 601-610).

Siau, K., Wand, Y., & Benbasat, I. (1997). The relative importance of structural constraints and surface semantics in information modeling. *Information Systems, 22*(2-3), 155-170.

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

Smith, G. F. (1988). Towards a Heuristic Theory of Problem Structuring. *Management Science*, *34*(12), 1489-1506.

Srinivasan, A., & Te'eni, D. (1995). Modeling as constrained problem solving: An empirical study of the data modeling process. *Management Science*, *41*(3), 419-434.

Tarafdar, M., Tu, Q., & Ragunathan, T. S. (2010). Impact of technostress on end-user satisfaction and performance. Journal *of Management Information Systems, 27*(3), 303-334.

Todd, P., & Benbasat, I. (1987). Process tracing methods in decision support systems research: Exploring the black box. *MIS Quarterly*, *11*(4), 493-512.

Vessey, I., & Conger, S. (1993). Learning to specify information requirements: The relationship between application and methodology. *Journal of Management Information Systems, 10*(2), 177-201.

Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 261-285). Hillsdale, NJ: Lawrence Erlbaum Associates.

Voss, J. F., Vesonder, G. T., & Spilich, G. J. (1980). Text generation and recall by high-knowledge and low-knowledge individuals. *Journal of Verbal Learning and Verbal Behavior, 19*, 651-667.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, *12*(4), 5-34.

Wiley, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory and Cognition, 26*(4), 716-730.

Witt, G. C., & Simsion, G. C. (2004). *Data modeling essential: Analysis, design and innovation* (3rd ed.). San Francisco, CA: Morgan Kaufmann.

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology, 43*(1), 111-122.

Appendix A: Experimental Materials

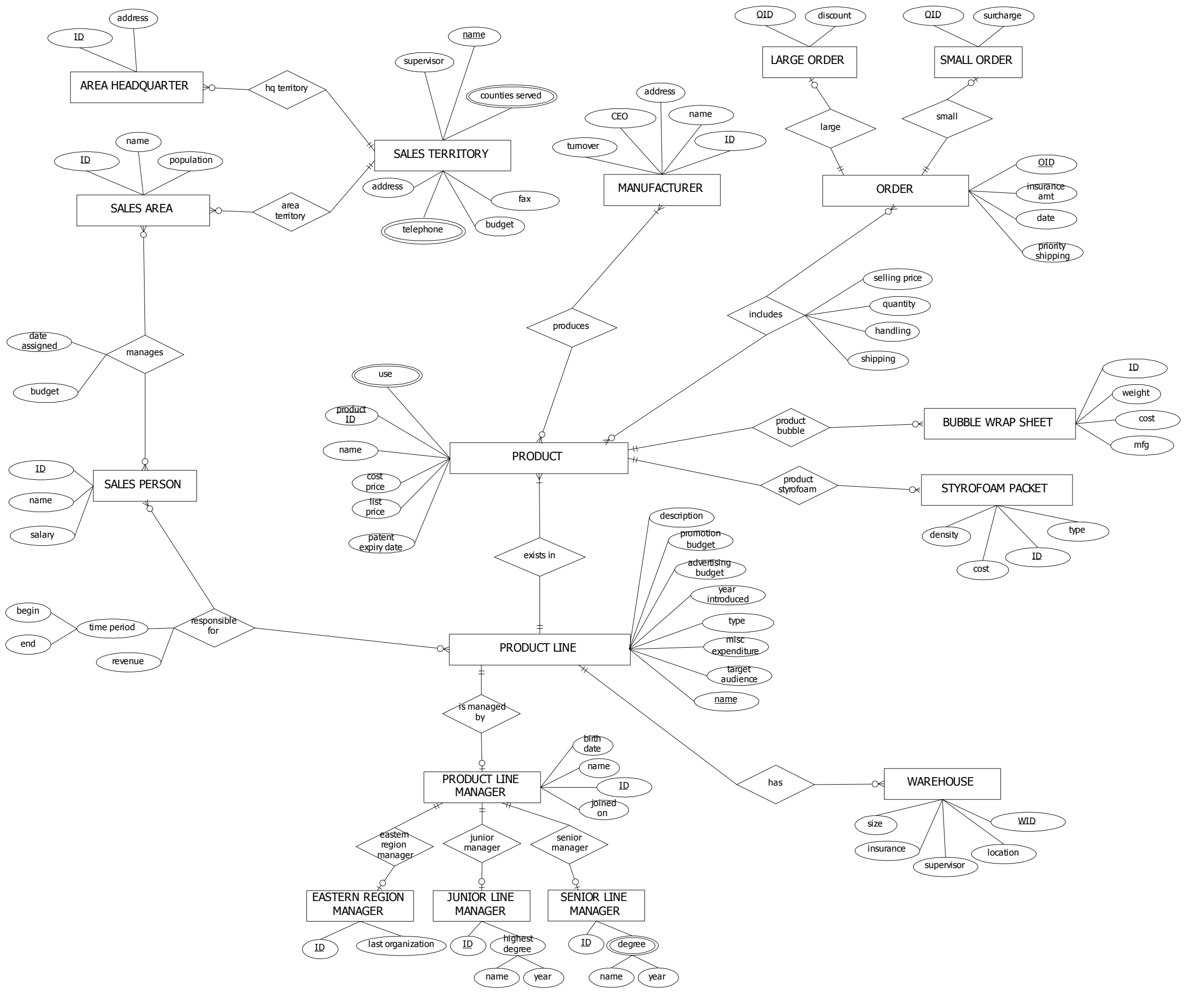


Figure A1. ER Sales Schema (Familiar Application Domain)

Table A1. Excerpt from the Data Dictionary for the Sales Schema

Entity type/relationship and attribute	Description
SALES TERRITORY: Sales region created by geography	
Name	Name of the sales territory
Supervisor	Name of the supervisor for the sales territory
Fax	Fax number
Counties served	Counties that are served by the sales territory
Budget	This year's budget for the sales territory
Telephone	Telephone(s)
Address	Address of the sales territory

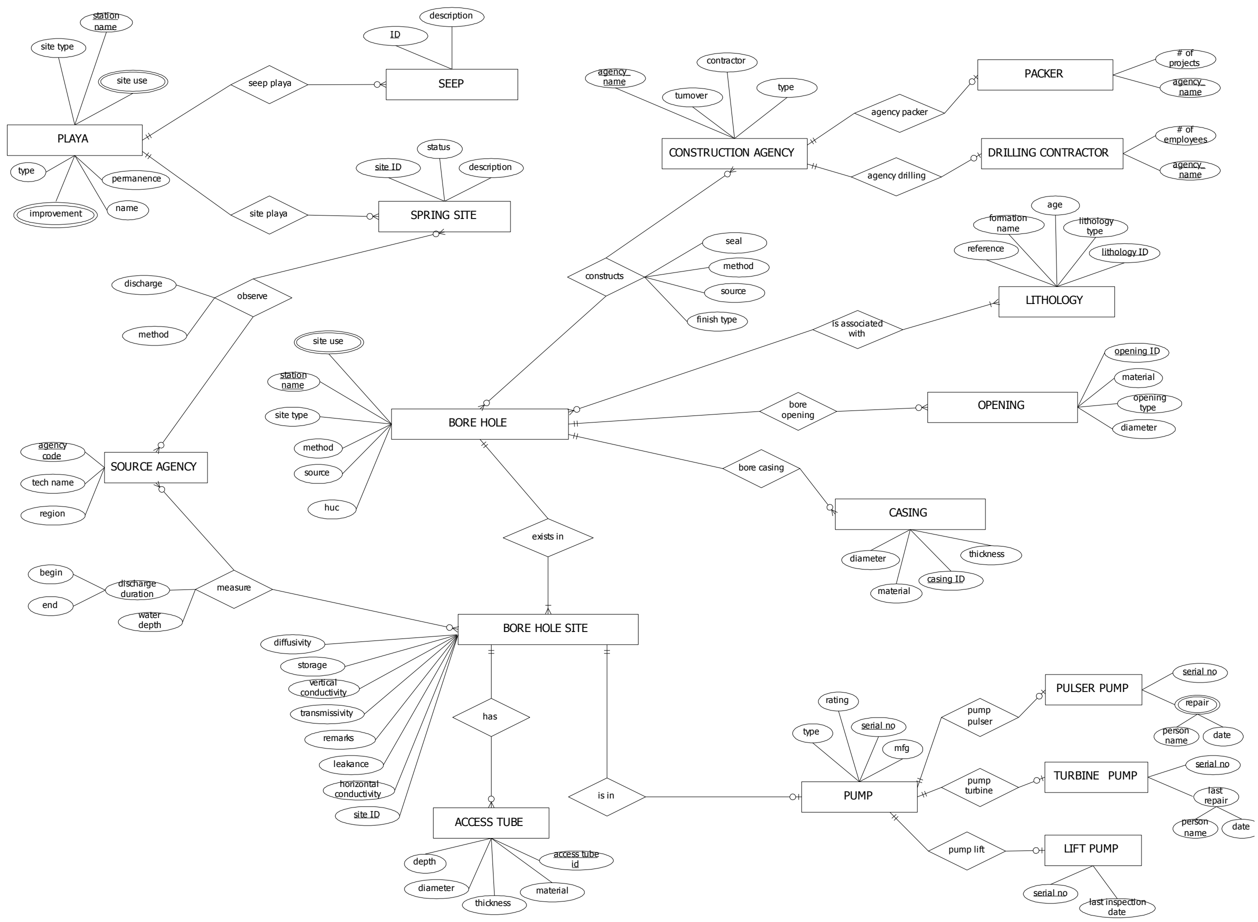


Figure A2. ER Hydrology Schema (Unfamiliar Application Domain)

Table A2. Excerpt from the Data Dictionary for the Hydrology Schema

Entity type/relationship and attribute	Description
PLAYA: A concentrated discharge of ground water to the surface that has an extent (i.e., area)	
Station name	Name of the surface water station
Site use	Purpose(s) for which site was constructed
Site type	Type of site, e.g., stream, lake or reservoir, estuary
Permanence	Permanence of discharge at spring
Name	Name by which the spring is known locally
Improvement	Type of improvements constructed at or in association with spring
Type	Type of spring (e.g., artesian, fracture, geyser, perched)

Appendix B: Participants' Background

Participant	Total years of work experience	Total years of database-related work experience	Highest earned qualification	Age	Gender
H-1[1]	Less than 2 years	Less than 2 years	High-school diploma	20-25 years	Male
H-2	Between 2 and 5 years	None	High-school diploma	20-25 years	Female
H-3	Less than 2 years	Less than 2 years	High-school diploma	20-25 years	Male
H-4	Less than 2 years	Less than 2 years	High-school diploma	20-25 years	Male
H-5	Between 2 and 5 years	Less than 2 years	High-school diploma	20-25 years	Male
H-6	Between 5 and 10 years	Less than 2 years	High-school diploma	20-25 years	Female
L-1	Between 2 and 5 years	None	High-school diploma	20-25 years	Male
L-2	None	None	High-school diploma	20-25 years	Female
L-3	Between 2 and 5 years	Less than 2 years	High-school diploma	20-25 years	Male
L-4	Between 2 and 5 years	Less than 2 years	High-school diploma	20-25 years	Female
L-5	Between 5 and 10 years	Less than 2 years	High-school diploma	20-25 years	Male
L-6	Between 2 and 5 years	None	High-school diploma	20-25 years	Female

[1]We explain this designation in the Section 3.3.1.

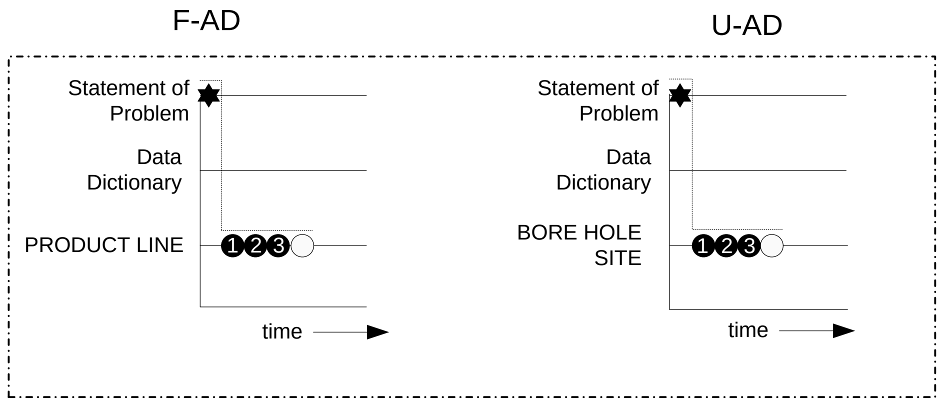
Appendix C: Conceptual Schema Comprehension Tasks

The simplest conceptual schema understanding tasks are known as conceptual *schema comprehension tasks*. Such tasks, which researchers have researched quite widely, may be either syntactic or semantic in nature (see, e.g., Khatri et al., 2006; Kim & March, 1995; Siau, Wang, & Benbasat, 1997). Khatri et al. (2006), for example, use the following examples of syntactic and semantic comprehension tasks, respectively.

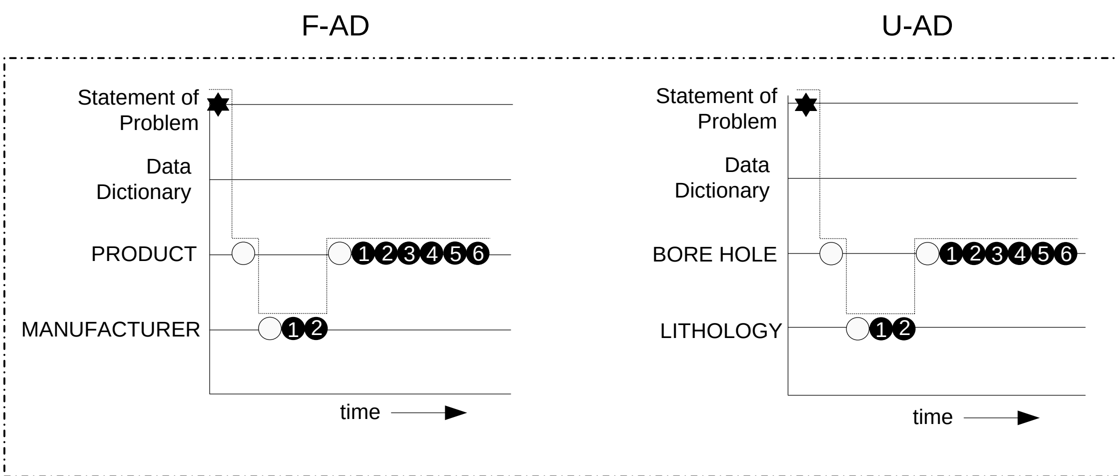
> What is the minimum:maximum cardinality of the relationship between PRODUCT and ORDER? (a) 0:M and 0:M; (b) 1:M and 0:M; (c) 0:M and 1:M; (d) 1:1 and 0:M
>
> A SALES PERSON is responsible for (a) exactly one PRODUCT LINE; (b) at the most one PRODUCT LINE; (c) no more than one PRODUCT LINE; (d) zero or more PRODUCT LINEs

These two types of schema comprehension tasks focus on the notations directly available in the schema (ISDK). As such, one may solve both syntactic and semantic comprehension tasks by referencing the conceptual schema directly. Therefore, one can solve such tasks efficiently and effectively using ISDK alone. In other words, knowledge of the application domain does not influence the solution of schema-comprehension tasks (Khatri et al., 2006).

Appendix D: Parsimonious Solutions for Schema-based Problem-solving Tasks



Task 1a



Task 1b

Figure D1. Parsimonious Solutions for Schema-based Problem-solving Tasks

Appendix E: Descriptive Response to Tasks

Solutions for Schema-based Problem-solving Tasks

Table E1. Descriptive Data for Solutions to Task 1a

	Familiar application domain (F-AD)		Unfamiliar application domain (U-AD)	
Participant	References to data dictionary	Search for E/Rs (# attributes referenced)	References to data dictionary	Search for E/Rs (# attributes referenced)
H-1	–	Product line (4)	–	Bore hole site (3), Bore hole
H-2	–	Product line (3), Product	–	Bore hole site (3), Bore hole
H-3	–	Product line (3)	–	Bore hole site (3)
H-4	–	Product line (3)	–	Bore hole site (3), Bore hole (1)
H-5	–	Product line (3), Product	–	Bore hole site (3)
H-6	–	Product line (3)	–	Bore hole site (3)
L-1	–	Product line (3), Product	–	Bore hole site (3), Bore hole
L-2	–	Product line (3)	–	Bore hole site (2), Bore hole (1)
L-3	–	Product line (3), Product	–	Bore hole site (3), Bore hole
L-4	Product line (target audience)	Product line (3)	–	Bore hole site (2), Drilling contractor (1), Constructs (3), Casing (1), Opening (1), Pump
L-5	–	Product line (3)	Borehole site (site ID)	Bore hole site (4), Borehole (1)
L-6	–	Product line (1), Sales person, Sales area, Sales Territory	–	Bore hole site (3)

Table E2. Descriptive Data for Solutions to Task 1b

	Familiar application domain (F-AD)		Unfamiliar application domain (U-AD)	
Participant	References to data dictionary	Search for E/Rs (# attributes referenced)	References to data dictionary	Search for E/Rs (# attributes referenced)
H-1	Product (all attributes)	Product, Manufacturer (3), Product Line (1), Order (1)	Borehole (source)	Lithology (2), Bore hole (1), Opening, Casing, Construction agency, Constructs (4)
H-2	–	Product (1), Manufacturer	Borehole	Bore hole (5), Constructs (4)
H-3	–	Manufacturer (4), Product (1)	–	Lithology (2), Bore hole (3), Opening (1), Bore hole site, Construction agency
H-4	–	Manufacturer (3)	Borehole (HUC)	Bore hole (1), Lithology (2)
H-5	Product	Manufacturer (1), Product (1), Product Line (1)	Lithology Borehole	Constructs (1), Borehole (1), Lithology
H-6	–	Product, Manufacturer (3)	–	Lithology (2), Constructs, Bore hole
L-1	Order (date)	Order (1), Product, Manufacturer (1)	–	Lithology (5), Bore hole (6), Casing, Opening

Table E2. Descriptive Data for Solutions to Task 1b

L-2	–	Manufacturer (2), Product	–	Lithology (2), Bore hole, Construction agency, constructs
L-3	–	Manufacturer (2), Includes (3), Order (4), Small Order (1), Large Order (1), Product (6)	Lithology Borehole (source, HUC) Constructs	Lithology (2), Bore hole site, Bore hole (3), Constructs, Construction agency
L-4	–	Manufacturer (5), Product (6), Product Line (7)	–	Construction agency, Constructs (1), Source agency, Opening
L-5	–	Manufacturer (3), Order, Product, Product Line (1)	Borehole (HUC, source)	Lithology (3), Bore hole (6)
L-6	–	Manufacturer (1), Product (1), Product Line (1)	–	Bore hole, Constructs, Lithology (2)

Solutions for Inferential Problem-Solving Tasks

Table E3. Descriptive Data for Solutions to Task 2a

	Familiar application domain (F-AD)			Unfamiliar application domain (U-AD)		
Participant	**References to data dictionary**	**Search for E/Rs (# attributes referenced)**	**References other than the schema and data dictionary**	**References to data dictionary**	**Search for E/Rs (# attributes referenced)**	**References other than the schema and data dictionary**
H-1	-	Product (1), Manufacturer	1) One reason would be that's not their forte… 2) …they don't make everything. So if you think of yourself as a retailer or grocery store the company that makes cereal that you sell probably does not make Kleenex or toilet paper or movies or candy that you sell may not have capacity to make certain things… 3)...they might not be the most significant manufacturer of that product 4)…you can buy it cheaper or better quality from somewhere else you don't ask company a to make it you only ask company b, so company a would not be listed as the manufacturer 5) might be because of copyright… 6) …they may not have the know-how to do even though they have the capacity or the right to make it 7) it might be a hi tech product you would purchase it from some that does make it.	Lithology	Bore hole, Bore hole site, Lithology, is associated with	1) …of course there's possible error where someone neglected to include it but that's not purposely done.

Table E3. Descriptive Data for Solutions to Task 2a

H-2	–	Manufacturer, Product (1), Product line	1) …manufacturers tend to segment themselves…let us say just even a stuffed animal there is a good chance that you are not going to be manufacturing heavy machinery… 2) it depends on your manufacturing capabilities 3) the machinery that you have 4) …and generally just from a business concept it's very difficult to remain profitable when you spread yourself out….and different segments of the market…hard…to manufacture everything 5) …may not necessarily want to manufacture a product	Bore hole, Lithology	Bore hole, Lithology (1), is associated with	1) Well, obviously if lithology only has to do with rock or rock formation, then there wouldn't necessarily have to be any rock around a well in the ground if there isn't any rock around. Then obviously you can't have any information about lithology 2) …the borehole wasn't necessary anymore. It was maybe filled back in.
H-3	–	Manufacturer, Product, produce	1) manufacturer do not have the capability to manufacture… 2) they don't have the right equipment, the right expertise 3)…products are outdated 4) …don't manufacture those anymore cause they don't think that they are cutting edge 5) …making the money that they want to make, 6)…they are at their capacity... from a capacity standpoint so they just stick with again what makes them the most money or what's the easiest to produce 7)…don't have the resources to make certain products	–	Lithology, Bore hole, is associated with, Borehole site	1) …maybe you have some extra ones on hand just in case the ones you have end up not working out so they haven't been used yet just because they're backups. 2) …maybe you bought a new one and haven't got it out to the site yet…
H-4	–	Produces, Manufacturer (1)	-	Lithology	Lithology, Bore hole	1) I don't know I really don't know a lot…
H-5	-	Manufacturer, produces, Product	1) …[product] was discontinued 2) another manufacturer took over production 3) …[manufacturer] went out of business 4) …another manufacturer is already manufacturing that one they might not be manufacturing it	-	Lithology, Bore hole, is associated with	-

Table E3. Descriptive Data for Solutions to Task 2a

H-6	–	Product (1), Manufacturer, produce	1) …wanted to focus on a particular set particular types of product 2) …no one…can…produce everything; lack capability 3) …they lack the capacity to produce a given product 4) …they lack the capability meaning machinery, skilled labor etc… 5) cost too much to produce; cost too much 6) …it is not profitable enough; product is not profitable enough 7) no one want to buy it	Bore hole, Lithology	Lithology	1) …they may have just not encountered it, not encountered it yet… 2) …just those that they found already yeah so basically just they don't apply…
L-1	-	Manufacturer, Product, produces	1) …based on what I know it would make sense because manufacturers don't manufacture everything out there, that's why there are many manufacturers in the world 2) …they don't have to manufacture any products…would make sense they wouldn't manufacture certain products	–	Bore hole, Lithology, is associated with	1) …it's allowed which is what lead to the situation Im sure there's a reason why but I know it…
L-2	–	Manufacturer, produces, Product	1) …that we used to sell before but we no more sell this product 2) manufacturer doesn't need to produce that product 3) …they might produce some product in the future.	Lithology, Bore hole	Lithology (4), Bore hole (5), is associated with	-
L-3	–	Manufacturer, produces, Product	1) there is nothing requiring them to… 2) and furthermore a manufacturer can exist without producing any product	-	Bore hole, Lithology, is associated with	-
L-4	–	Product (3), Manufacturer (1), produce	1) …maybe the licensing issued… 2) maybe the manufacturer change the place so it's not really close to the company, so the company have to change another manufacturer 3) even maybe they change the CEO [and] he or she didn't like this company maybe just disconnected, discontinued	Lithology	Bore hole, Lithology, is associated with	-

Table E3. Descriptive Data for Solutions to Task 2a

L-5	–	Manufacturer, Product (1), includes (2)	1) …each manufacturer is going to have its own equipment, different equipment from some of the other manufacturers… 2) …they may specialize or have specific technology related to manufacture 3) cost they get from their suppliers… 4) …manufacturers have their own equipment or different equipment from one place to the next and they probably have special equipment dedicated to male one or more of those products	Lithology	Lithology, Bore hole	-
L-6	–	Manufacturer, Product, produce	1) it is just not profitable… 2) maybe they specialize in certain products with high return…; they are probably specializing in specific products 3) …maybe they don't have the facilities to make those certain products 4) maybe the products are not popular anymore so they don't make them anymore…	–	Lithology (1), Bore hole	-

Table E4. Descriptive Data for Solutions to Task 2b

	Familiar application domain (F-AD)			Unfamiliar application domain (U-AD)		
Particip ant	References to data dictionary	Search for E/Rs (# attributes referenced)	References other than the schema and data dictionary	References to data dictionary	Search for E/Rs (# attributes referenced)	References other than the schema and data dictionary
H-1	-	Sales territory (3), Sales area (1), manages (1), sales person (1)	1) contact for the sales territory…. 2) …past sales data for that area 3) …possibly contact information for previous sales manager who used to manage that area… 4) …what modes of travel he would want to go through… 5) …number of customers… 6) …he would want to know the past sales data 7) …and maybe why past customers no longer purchase from the company, like if a customer was thinking of suing a company he would know not to call on that customer any more 8)…he would want to know his salary/commission and how he would be paid for this.	Discharge	Spring site (2), Playa (5), Source agency, observe (2)	-

Table E4. Descriptive Data for Solutions to Task 2b

H-2	–	Sales person, Sales area (1), Sales territory, manages (1)	1) …not all sales people need to be managers… 2) if you weren't given control over certain sales territory then obviously you can't take control of certain sales areas… 3)…also it will tell you the kind of approaches you will need for your management techniques… 4) …and maybe even tell you how many subordinates you'll need to hire if you have that power to do so for each sales area…	Source agency, Spring site	Spring site (2), Source agency (1), observe (2)	1) I don't know if there is any jurisdiction that a source agency will have over certain regions…. 2) …they will have to take into account when they observe it 3) …obviously if environmental conditions aren't going to be conducive to be able to observe certain things about the site…
H-3	–	Sales area (3), Sales person, manages (2), Product line (1), Warehouse	1) know how quickly they can ship… 2) they need to know just overall the number of customers that are going to be in that area…	–	Spring site (3), observe (2), Playa (1)	-
H-4	–	Sales area (1), manages (2)	-	observe (discharge, method)	observe (2)	-
H-5	-	Sales person, manages (2), Sales area (3), Sales territory (3), Area headquarter	1) so if he's in charge of the sales area, he's most likely just going to need to be concerned with that area…	observe (method)	Spring site (3), observe (2), Playa	1)…we need to make sure we record how that measurement was taken and the actual measurement…
H-6	–	Sales area (3), Sales territory (7), manages (1)	1) everybody looks at their bottomline	-	Lithology	-
L-1	-	Sales person, Sales area (2), manages (2)	-	–	Spring site (3), observe (2)	-
L-2	–	Sales person, manages (1), Sales area (3), Sales territory, Area headquarter	-	-	Source agency, observe (2), Spring site (2), Playa (3)	-
L-3	–	Sales person, manages, Sales area	-	observe (discharge, method)	Source agency, Spring site (2), observe (2)	-
L-4	–	Sales area (1), Sales person (1), Sales territory (3), Product line manager, manage (1)	-	Source agency (region, tech name), observe (discharge, method)	Source agency (3), Spring site (1), measure (2), observe (1), Bore hole site (2)	-

Table E4. Descriptive Data for Solutions to Task 2b

L-5	–	Sales area (1), Product line, Sales person, Sales territory	1) …family sizes… 2) …income level of the sales area… 3) the number of kids… 4) …the average age in the county 5) size of the sales area would be important… 6) …number of computers in a household 7) …telephone lines…	Source agency, Spring site	Source agency, Spring site (1)	1) Other factors they are going to have to include are weather, the time of the day, time of year… 2) …amount of rainfall of the year or the month or the week, depending on how specific they want the information 3) I guess population of the area, maybe the amount of traffic that location receives as far as human traffic, wild life in the area, soil type you could even go into plant types bush types…
L-6	–	Sales area (2), sales territory (5), Area headquarter	-	–	Source agency (1), Spring site (2), observe (2), Playa (6)	-

Appendix F: Detailed Transition Data

Table F1. Number of Problem-solving Transitions on Schema-based Problem-solving Tasks (Tasks 1a and 1b)

	Participants	F-AD	U-AD
Task 1a			
H-ISDK participants	H-1	1	3
	H-2	4	7
	H-3	1	1
	H-4	3	5
	H-5	9	3
	H-6	1	3
	Median	**2**	**3**
L-ISDK participants	L-1	3	6
	L-2	1	4
	L-3	3	4
	L-4	5	7
	L-5	1	4
	L-6	7	1
	Median	**3**	**4**
Task 1b			
H-ISDK participants	H-1	7	11
	H-2	4	5
	H-3	6	13
	H-4	1	6
	H-5	4	7
	H-6	3	5
	Median	4	6.5
L-ISDK participants	L-1	9	11
	L-2	6	9
	L-3	9	22
	L-4	5	4
	L-5	8	9
	L-6	4	5
	Median	**7**	**9**

Table F2. Number of Problem-solving Transitions on Inferential Problem-solving Tasks (Tasks 2a and 2b)

	Participants	F-AD	U-AD
Task 2a			
H-ISDK participants	H-1	16	11
	H-2	13	12
	H-3	17	8
	H-4	2	5
	H-5	16	15

Table F2. Number of Problem-solving Transitions on Inferential Problem-solving Tasks (Tasks 2a and 2b)

	H-6	18	7
	Median	**16**	**9.5**
L-ISDK participants	L-1	8	7
	L-2	18	17
	L-3	13	6
	L-4	11	10
	L-5	12	7
	L-6	13	5
	Median	**12.5**	**7**
Task 2b			
H-ISDK participants	H-1	21	8
	H-2	18	17
	H-3	12	10
	H-4	5	4
	H-5	16	10
	H-6	8	7
	Median	**14**	**9**
L-ISDK participants	L-1	7	2
	L-2	14	7
	L-3	10	8
	L-4	12	10
	L-5	14	10
	L-6	4	8
	Median	**11**	**8**

Table F3. Number of References Not Directly Related to Experimental Material on Inferential Problem-solving Tasks (Tasks 2a and 2b)

	Participants	**F-AD**	**U-AD**
Task 2a			
H-ISDK participants	H-1	6	1
	H-2	4	2
	H-3	7	2
	H-4	0	0
	H-5	4	0
	H-6	7	2
	Median	**5**	**1.5**
L-ISDK participants	L-1	2	0
	L-2	1	0
	L-3	1	0
	L-4	3	1
	L-5	4	0
	L-6	4	0

Table F3. Number of References Not Directly Related to Experimental Material on Inferential Problem-solving Tasks (Tasks 2a and 2b)

	Median	**2.5**	**0**
Task 2b			
H-ISDK participants	H-1	8	0
	H-2	4	3
	H-3	2	0
	H-4	0	0
	H-5	1	1
	H-6	1	0
	Median	**1.5**	**0**
L-ISDK participants	L-1	0	0
	L-2	0	0
	L-3	0	0
	L-4	0	0
	L-5	6	3
	L-6	0	0
	Median	**0**	**0**

Appendix G: Summary of Participant Numbers in IS Protocol Analysis Studies

Research	Number of participants
Vessey and Conger (1993)	6
Srinivasan and Te'eni (1995)	14
Shaft and Vessey (1998)	24
Kim et al. (2000)	9+7
Purao et al. (2002)	2
Hungerford et al. (2004)	12
Burton-Jones and Meso (2006)	6

About the Authors

Vijay Khatri joined Indiana University in 2002, where he is an associate professor of information systems in the Operations and Decision Technologies department. He holds a B.E. degree (electronics) from Malaviya National Institute of Technology, a management degree from the University of Bombay, and a PhD degree from the University of Arizona. His research centers on issues related to data semantics, semiotics and conceptual database design, temporal databases, and data governance. More specifically, his research involves developing conceptual design techniques for management of data, especially for applications that need to organize data based on time and space. Vijay has published articles in journals such as *IEEE Transactions on Knowledge and Data Engineering*, *Information Systems*, *Annals of Mathematics and Artificial Intelligence*, *Information Systems Research*, *Journal of Management Information Systems*, *Decision Sciences Journal*, *IEEE Transactions on Software Engineering*, and *Communications of the ACM*. He serves on the editorial review board of the *Journal of Association of Information Systems*. He is a member of ACM and a senior member of IEEE.

Iris Vessey is Professor of Information Systems at the University of Queensland. She holds M.Sc., MBA, and Ph.D. Degrees from the University of Queensland, and has held positions at Pittsburgh University, The Pennsylvania State University, and Indiana University. Her research interests center on fit/alignment at both individual and organizational levels. Her research into individual problem solving focuses on the enduring notion of cognitive fit (developed in 1991) in system design and development. This research frequently encompasses the role of the application domain in systems cognition. Her research into organizational IS/IT alignment is based in the complexity theory worldview, which has significant implications for IS alignment that is sustainable, as well as for business agility. She has published in journals such as *Decision Sciences*, *Information Systems Research*, the *Journal of the Association for Information Systems*, the *Journal of Management Information Systems*, and *MIS Quarterly*. She is a Fellow of the AIS.