ABSTRACT

Title of Document: SENSEMAKING: CONCEPTUAL CHANGES, COGNITIVE MECHANISMS, AND STRUCTURAL REPRESENTATIONS. A QUALITATIVE USER STUDY.

Pengyi Zhang, Doctor of Philosophy, 2010

Directed By: Professor Dagobert Soergel
College of Information Studies

The purpose of this thesis is to improve our understanding of sensemaking process as a basis for building better systems to assist sensemaking. Sensemaking is the task of creating an understanding of a problem or task so that further actions may be taken in an informed manner. Sensemaking is a pre-requisite for many other tasks such as decision making and problem solving. An important part of sensemaking involves making clear the interrelated concepts and their relationships in a problem or task space.

This research investigated the question of how users create and use structured representations for sensemaking. It proposed and refined an iterative sensemaking model building upon previous sensemaking research, learning theories, cognitive psychology and task-based information seeking and use. In particular, the study focused on the processes, conceptual changes, and cognitive mechanisms used during users' sensemaking tasks. The qualitative, multi-case user study investigated how a

sample of fifteen users working with news writing and business analysis tasks structure their conceptual space with the assistance of note-taking and concept mapping tools. Data on the sensemaking process were collected from multiple sources including think-aloud protocols, screen movement recordings, interviews, and intermediate and final work products.

Using the iterative sensemaking model as an analytical and descriptive framework, the study captured the often idiosyncratic paths sensemakers took, ranging from planned, systematic to rather random, ad hoc patterns of "search--sensemaking" iterations. Findings also revealed various ways in which the iterations started and exited, which suggested that the heterogeneous patterns of sensemaking lie in the shifts from one iteration to the next, rather than in the iterations themselves. The knowledge structure was updated by accretion, tuning, and restructuring to produce the final knowledge representation and sensemaking product. Several cognitive mechanisms were used in processing new information, examining concepts and relationships, and examining anomalies and inconsistencies. They were used in bottom-up, top-down, and combined fashions to move the processes along and to trigger conceptual changes to the knowledge structure of users.

Based on these findings, the study argues that information system that aimed to assist sensemaking should provide an architecture that links structure, data, and sources that can be represented and manipulated in multiple formats. It should also provide integrated assistances at the task and cognitive mechanism levels.

The research contributes to sensemaking research by extending existing descriptive sensemaking models with an analytical framework that incorporates conceptual changes to the knowledge structure and cognitive mechanisms that trigger the processes and conceptual changes. Furthermore, the research identified core issues in designing information systems to assist sensemaking tasks and suggested design implications for sensemaking tools that may be useful in many settings such as learning, knowledge creation, organization, and sharing.

SENSEMAKING: CONCEPTUAL CHANGES, COGNITIVE MECHANISMS, AND STRUCTURAL REPRESENTATIONS. A QUALITATIVE USER STUDY

By

Pengyi Zhang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2010

Advisory Committee:
Professor Dagobert Soergel, Chair
Associate Professor Linda Aldoory, Dean's Representative
Associate Professor John Newhagen
Associate Professor Douglas W. Oard
Assistant Professor Yan Qu

Dedication

To my parents.

Acknowledgements

I am deeply indebted to my advisor and chair of my dissertation committee, Dr. Dagobert Soergel, who coached me through the PhD program and helped me in every possible way. I could not thank him enough for his dedication, support, and confidence in me. I also wish to thank members of my dissertation committee, Dr. Linda Aldoory, Dr. John Newhagen, Dr. Doug Oard, and Dr. Yan Qu, for their valuable suggestions and feedback throughout this research.

During my study at Maryland, I was fortunate to have worked with several faculty members. I thank them, especially Dr. Jimmy Lin, Dr. Doug Oard, Dr. Judith Klavans, and Dr. Yan Qu, for giving the opportunity to work on their projects.

This dissertation was partially supported by the Eugene Garfield Doctoral Dissertation Fellowship. I wish to express my appreciation to Beta Phi Mu and Dr. Garfield for this generous award.

I am thankful to the students who took their time participating in this research and sharing their experience and insights with me.

I was lucky to have several wonderful friends at Maryland whom I received much help from and shared a lot of fun with. I wish to thank Xiaohan Hu, Xiaoxue He, Wei Jin, Lin Yuan, and Lynne Plettenberg for sharing this journey with me.

My greatest appreciation goes to my husband, Hongzhong Zhou. Without his unconditional love, continuous support, and tremendous help, I could have not made this accomplishment. Our lovely daughter, Momo, came along while I was working on this research and has brought so much joy ever since. Words cannot express how much I love you.

Table of Contents

List of Tables

List of Figures

List of Illustrations

Chapter 1: Introduction

This thesis aims to advance our understanding of individual sensemaking. It does so by developing and refining a sensemaking model from empirical user studies. It provides a design framework for information systems and tools that assist users to search for, create, and organize a structured conceptual space, as a basis for performing a task or making a decision.

Sensemaking is a pre-requisite for many other tasks such as decision making and problem solving. An important part of sensemaking involves making clear the interrelated concepts and their relationships in a problem or task space.

1.1 Problem Statement

1.1.1 What is Sensemaking?

People encounter sensemaking tasks every day. Sensemaking often involves the following steps:

- Recognize a knowledge gap
- Possibly generate an initial structure or model of the knowledge needed to complete the task – concepts, relationships, and hypotheses
- Search for information
- Analyze and synthesize information to create an understanding
- Create a task product based on this understanding in the form of a report, decision, or solution to a problem

Typical examples of sensemaking are shown in Box 1-1:

Example 1

A business analyst gathers, analyzes, and uses information about a product and its alternatives, customers, and competitors, and makes planning suggestions for the organization. The task is familiar to the analyst, but the domain or product may be new from time to time. The analyst needs to extract from the information found the related entities and concepts, the relationships between concepts and entities, create an understanding of the various relationships, and make reasonable suggestions based on the identified patterns and relationships.

Example 2

A patient has been diagnosed as having high blood pressure. He needs to learn about the condition. The problem and the domain are both new to this patient. He needs to find out about the causes, consequences, treatment options, influences, and so on, of the condition. The major task is for him to develop a mental model of the condition and relate it to his previous knowledge structure so that he can make decisions or take action.

Box 1-1: Examples of Sensemaking

Sensemaking is a very common activity. In fact, information use can be viewed as a continuum of various levels of sensemaking, ranging from fitting information directly to its need (such as a task like catching a plane that can be supported with a factoid answer – the departure time) to very complex sensemaking activities that require synthesis and assimilation of information into the user's existing knowledge structure to establish an understanding represented in certain ways that further actions may be based upon. Sensemaking is particularly important on this end of information use when people are faced with new situations and less structured tasks.

People make sense of situations or problems in various ways. For example, the patient may talk to doctors, nurses, even friends to learn about high blood pressure. He may also try to learn about the condition by reading medical essays in magazines and searching the Web. In organizational settings, a person may start her sensemaking by observing how others behave and listening to what they say. Many people use information systems as sensemaking tools. With the advance of technologies in information retrieval (IR), standard IR systems can support reasonably well the search for pieces of relevant information when the user can identify her information need or knowledge gap to a certain extent. Research has also started to develop systems to support exploratory search, which is very important to sensemaking.

But searching is only one part of the sensemaking process. Beyond search, sensemaking is about understanding complex things and their relationships, i.e. building a rich representation from a large amount of data. Most information systems, including search engines, are designed to help users to get the information they may find useful for a task. Often the retrieved information is either too much or too fragmented and the relationships are obscure. In order to make use of the information they found, users need to understand the pieces of interrelated information they found, identify patterns, and build on their previous knowledge to create an updated understanding. Most sensemaking tasks are characterized by the search for and creation of a structured representation of the situation, problem, or domain. What users need are sensemaking tools that facilitate the creation of such conceptual structures. It is important to achieve better understanding of user processes in sensemaking tasks in order to design better tools to assist sensemaking. This is the main motivation for the study.

Previous research on sensemaking has examined the searching aspects extensively, while the work on the construction of representations has been by and large descriptive and leaves unanswered several important questions about the conceptual changes of the representations and the cognitive processes and mechanisms. Research in education (especially learning theories), cognitive psychology (cognitive processes and structures), and information seeking and use behaviors all bring useful insight to sensemaking research. This thesis intends to develop a sensemaking model that provides explanatory power by examining other areas, including cognition and learning, and by studying users' sensemaking processes with the assistance of information systems.

The advance in information technologies such as information retrieval and visualization provides an opportunity for examining the use of such technologies for sensemaking, which was not possible earlier. Sensemakers extract main concepts and relationships from raw material to establish an understanding. When a sensemaking task is difficult, sensemakers usually employ external representations to store the information for repeated manipulation and visualization (Stefik, Baldonado et al. 1999). With the advance of technology in information extraction, information retrieval, and information visualization, sensemaking tools could incorporate such automated techniques to extract concepts and relationships, to help users with identifying patterns in the fragmented pieces of useful information from various sources, and to provide intuitive pictures of the mental models and processes of users.

1.1.2 Implications

This thesis aims to investigate how users create and use structured representations for sensemaking. Sensemaking is sometimes a challenging task for information users. Assisting users with their sensemaking is a challenging task for systems. Automatic processing is not advanced enough to achieve human insights and inferences with regards to user context. However, systems can assist users with processing massive data through the following techniques:

- Extracting main entities or concepts
- Presenting results in a way that patterns may be recognized more easily
- Aligning system outputs with user task outputs

Sensemaking tools are useful in many settings. They can be used as:

- Knowledge management and sharing tools by individuals or small groups for collective sensemaking
- Learning and reflection tools for educational uses which reveal and record students' learning processes and outcomes
- Devices for eliciting and modeling expert knowledge when used by experts on certain tasks

The "outcome" (Dervin 1980; Dervin 1992; Dervin 1998) of successful sensemaking is very important for accomplishing tasks, and may be needed for future reference and use. In addition to the final product of sensemaking (often embodied in some form of a formal report, a decision to guide certain actions, or a solution to a problem), intermediate products of sensemaking are also very important and may be worth recording and sharing.

Helping users retrieve the right information is only half the battle; assisting users with making sense of what they found is the next frontier in information system design. This study will contribute to our understanding of sensemaking processes, and thereby give a better foundation for design of systems and tools that support sensemaking.

1.2 Research Question

The overall research question is:

How do users make sense of a complex situation (i.e., structure their conceptual spaces about a topic, task or problem) with the assistance of sensemaking tools?

This broad question touches on issues related not only to process but also to changes to the knowledge structure in the conceptual space, and cognitive mechanisms that drive the process and conceptual changes. Related foreshadowing questions address these specific aspects:

1. What process do users go through in structuring their conceptual space?
2. How do users' conceptual models evolve? I.e. what conceptual changes happen to users' conceptual models, and how?
3. What cognitive mechanisms do users use, and how?
4. What information do users search for and use to create the conceptual models?

5. What formats and system functions do users use to create and represent their conceptual models?

6. How do users organize raw information/data, conceptual maps, intermediate task outputs, and the final product of sensemaking such as a news story?

1.3 Contributions

1 **A new sensemaking model.** The main contribution is the iterative sensemaking model proposed and examined, building upon previous sensemaking research, learning theories, cognitive psychology and task-based information seeking and use. The model provides a descriptive and analytical framework in examining sensemaking process which may be applied to settings beyond the user cases examined in this research.

2 **Sequence diagram for visualizing and analyzing the sensemaking process and activities**.

3 **Recurring modules of the sensemaking process.** Research in sensemaking has discovered the idiosyncratic and iterative nature of sensemaking; this dissertation reveals the recurring modules of the sensemaking process – various ways in which the sensemaking iterations started and ended. Findings suggested that the heterogeneous patterns of sensemaking lie in the shifts from one iteration to the next, rather than in the iterations themselves.

4 **Understanding of conceptual changes and cognitive mechanisms.** Guided by the model, the results lead to a better understanding of conceptual changes that occur in the sensemakers' knowledge structure and the cognitive mechanisms used to trigger these changes during the sensemaking process. These results deepen the understanding of information behavior, task-oriented learning, and eventually organizational learning.

5 **Recommendations for the design of sensemaking support systems.** The design framework and recommendations for design functionality based on the case studies provide the basis for improved sensemaking support systems.

Chapter 2: Literature Review and Theoretical Framework

This chapter reviews related literature and proposes an integrated sensemaking model based on the literature. Four areas contribute to the theoretical framework and inform the design of the sensemaking tool:

- Sensemaking and task-based information seeking
- Cognitive processes and structures
- Learning theories
- Design of related tools, including knowledge representation, information extraction, concept mapping, and other tools

Sensemaking research has identified several processes that are involved in users' sensemaking activities while performing various tasks. Several sensemaking models were proposed for different purposes, such as

- to describe the sensemaking processes, either of particular or generic user groups (Krizan 1999; Pirolli and Card 2005; Qu and Furnas 2007);
- to provide an analytical abstraction derived from empirical user studies (for example, Russell, Stefik et al. 1993);
- to guide the sensemaking practice of certain groups of sensemakers, for example decision makers (Kurtz and Snowden 2003).

These models together provide insights and a framework the model proposed in this thesis as presented in Section 2.5. However, the existing models are by and large descriptive in nature and do not explain:

- The cognitive processes and structures by which knowledge is created and stored during the sensemaking process
- How different types of conceptual changes happen during the sensemaking process
- What mechanisms trigger the changes and enable the assimilation of new information and the creation of a structural representation

This thesis aims to develop a sensemaking model that moves to a stronger basis for explaining sensemaking behaviors by examining other areas including cognition and learning theories.

Research in cognitive processes has examined:

- The various ways in which a conceptual structure may be changed (Chi 1992; Chi 2007)
- The mechanisms that provoke these changes (Toulmin, Rieke, & Janik, 1979; Kavale, 1980; Rumelhart & Norman, 1981; Vosniadou & Ortony, 1989; M. T. H. Chi, 1992; Johnson-Laird, 1999)
- The role of existing knowledge (Dole and Sinatra 1998)

Theoretical and empirical findings in these areas provide great insights to the conceptual changes happening during sensemaking processes. Researchers in cognition seem to suggest that:

- Knowledge is stored as connected concepts and relationships in the brain (Rumelhart and Ortony 1977; Carley and Palmquist 1992; Jonassen and Henning 1996).

- External representations (Zhang 1997; Zhang 2000) may be very useful in facilitating many cognitive tasks, including sensemaking.

To suggest design ideas for systems and tools to help with users' creation and representation of sensemaking products, this chapter also reviews design of related tools such as knowledge representation, concept mapping, information extraction, and other tools.

Research in knowledge representation suggests various ways in which intermediate and final products of sensemaking may be stored and manipulated, including concept maps, templates, outlines, and text representation. Concept mapping allows sensemakers to put raw and extracted information together in a meaningful representation and thus may assist the production of sensemaking outcomes. Information extraction techniques have been advanced enough to extract entities and relationships from text with satisfactory performance. Other tools, such as tools supporting exploratory search and analytical tools, are also reviewed.

Sections 2.1 to 2.4 review each of the above areas in detail, culminating in the definition of a new comprehensive sensemaking model in Section 2.5.

2.1 Sensemaking Models

Sensemaking is characterized as a series of continuing gap-defining and gap-bridging activities between situations (Dervin 1992; Dervin 1998). Several researchers have been studying the sensemaking processes and proposed descriptive models to illustrate them. This section reviews some of the representative models of sensemaking processes.

Sensemaking tasks often involve searching for information that is relevant for a task and then extracting and analyzing information to create an understanding on which to base decision or actions. Sometimes researchers refer to sensemaking strictly as the analysis, synthesis, and conceptualization part of the process, for example, the process of creating a representation and encoding data in that representation to answer task-specific questions (Russell, Stefik et al. 1993). However, since the searching activities and sensemaking activities are often closely intertwined, some researchers also refer to sensemaking as the overall process of creating an understanding, which includes both finding and understanding information (Pirolli and Card 2005).

Two types of distinctive activities emerge from the literature.

1. Seeking for information, followed by extracting and filtering the information found.
2. The iterative creation and updates of an understanding of the situation, especially connections (for example, people, places, and events) in order to anticipate their trajectories and act effectively (Klein, Moon et al. 2006).

Each category may include several cognitive and behavioral activities. They do not necessarily follow a linear, two-stage order. They may iterate several rounds until the goal is reached, intertwine with and influence each other, or spiral up to the final product.

2.1.1 Russell's Model

Russell and others (1993) did several case studies to explore the sensemaking process. They separate sensemaking from the target task, which could be a learning task, a decision making task, or a simple information processing task (Bystrom 2002; Hansen and Bystrom 2005).

Four main processes are involved in sensemaking, shown in Figure 2-1:

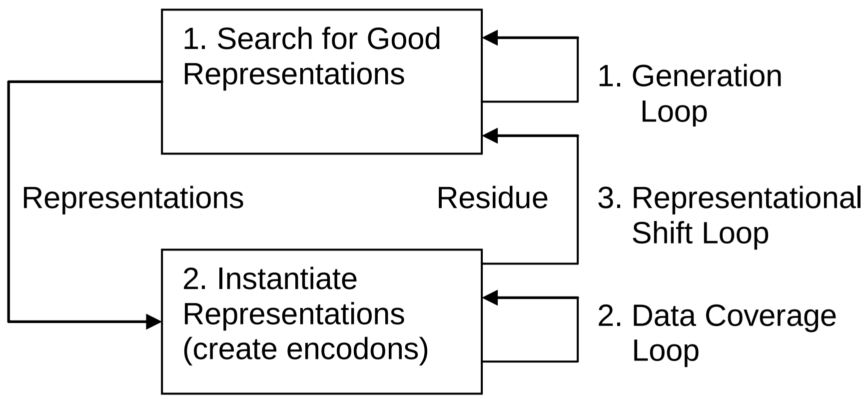


Figure 2-1: Russell's Sensemaking Model

1 **Search for representation** (generation loop): both representations and the procedures that use them are created.

2 **Create instances of representations** (data coverage loop): the sense maker identifies information of interest and encodes it in the representation.

3 **Modify representation** (representation shift loop): representations are modified during sensemaking when data is ill-fitted or missing.

4 **Consume instantiated schemas**: the sense maker uses the instantiated schemas in the task-specific information processing step. The schemas provide top-down or goal-directed guidance, by prescribing what to look for in the data, what questions to ask, and how the answers are to be organized.

Box 2-1: Russell's Sensemaking Model

Structural representation plays a crucial role in all processes. The generation loop represents the construction of a structural representation; the data coverage loop represents the fitting of data or evidence into the structure. When mismatch happens between the representation and the data, a residue of data that do not fit remains, and the representational shift loop takes place to reconstruct the representation.

Qu and Furnas (2007) further investigated the seeking of structural information in the sensemaking process. They separated the searching for structures from the searching for data in the sensemaking process. They also tried to integrate the two processes and emphasize the bi-directional relationship between search and representation construction. When the representation is inadequate, incomplete, or ill-formed, a need for changing, growing, or validating structures arises, shown in the following figure:

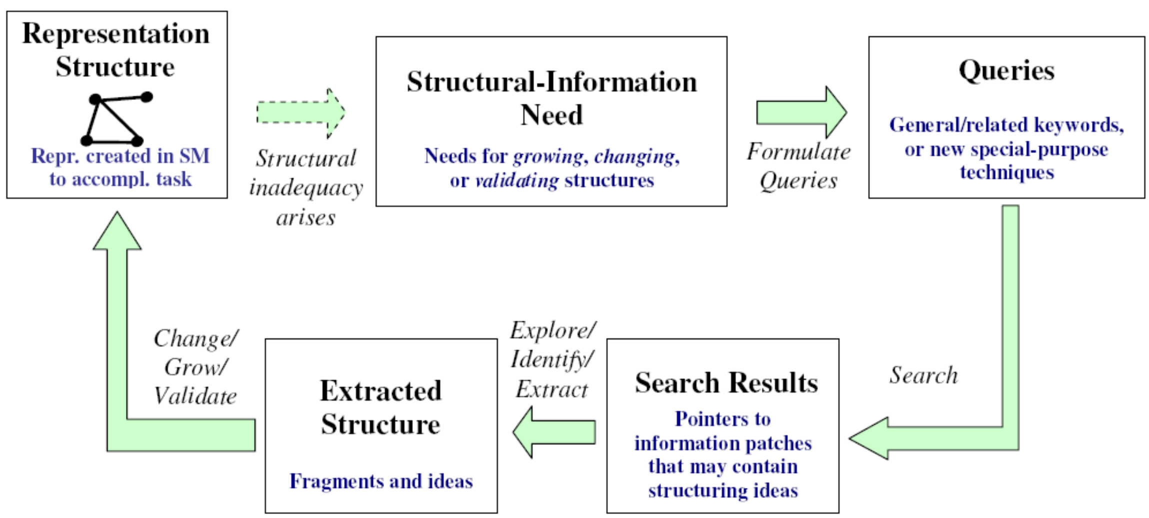


Figure 2-2: Structural Information Seeking (Qu and Furnas 2007)

2.1.2 The Cyclical Model of Intelligence Process

The process of intelligence creation and use, whether it is in government or business setting, follows a series of repeated and interrelated steps (Krizan 1999). Each step adds value to the inputs and together they create a substantially updated report. The analysis processes, i.e. sensemaking, convert information into intelligence for planners and decision makers.

The cyclical model of intelligence process is shown in Figure 2-3:

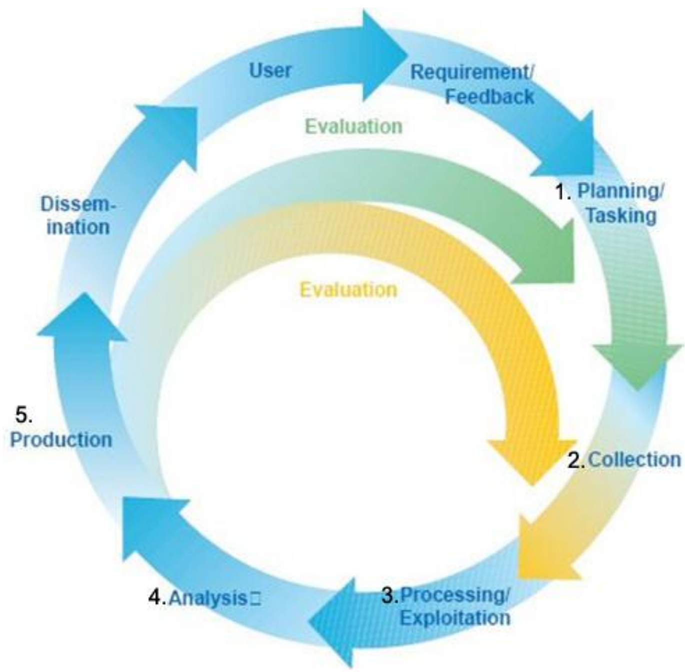


Figure 2-3: The Cyclical Model of the Intelligence Process (Krizan 1999)

In this model, five critical steps are identified, including planning/tasking, collection activities, processing of collected information, analysis, and production:

1. Intelligence needs are assigned or provided by customers to analysts. They are often complex and time-sensitive. Intelligence analysts need to interpret the customer requirements before the task can be

processed.

2. In the collection step, analysts acquire information from various sources, including people and information systems.

3. Processing is the selection of raw information based on its plausibility, expectablity, and support to intelligence issues.

4. In the analysis step, analysts try to make sense of the selected information and make higher level analysis including giving descriptions of the task domain, establishing explanations of phenomenon, interpreting cause and effects, and so on.

5. In the production step, the final product of intelligence process, “value-added actionable information”, is created. This means synthesizing all available sources, including the intermediate products of previous steps, to create a comprehensive assessment of an issue or situation.

The product of sensemaking, an intelligence report, is disseminated to the customers for evaluation and feedback, and the next round of sensemaking activities follows.

Box 2-2: The Cyclical Model of Intelligence Process

2.1.3 Pirolli & Card’s Model of Intelligence Analysis

Through cognitive task analysis, Pirolli and Card (2005) proposed a notional model of sensemaking, with two loops of activities:

- An information foraging loop that involves processes aimed at seeking information, searching and filtering it, and reading and extracting information into some schema
- A sensemaking loop that involves the iterative development of a mental model (a conceptualization) from the schema that best fits the evidence

There are ten processes and six presentations (ranging from external raw data to the final task presentation) of the sensemaking process for intelligence analysts. External data sources are raw evidence, largely in textual form. The "shoebox" is a much smaller subset of the external data that is relevant for processing. The evidence file refers to snippets extracted from items in the shoebox. Schemas are the re-representation or organized marshalling of the information so that it can be used more easily to draw conclusions. Hypotheses are the tentative representation of those conclusions with supporting arguments. Ultimately there is a presentation or other work product. The production of task output follows the path "Information → schema → insight → product."

The overall sensemaking process consists of information gathering, representation of the information in a schema that aids analysis, the development of insight through the manipulation of this representation, and the creation of some knowledge product or direct action based on the insight.

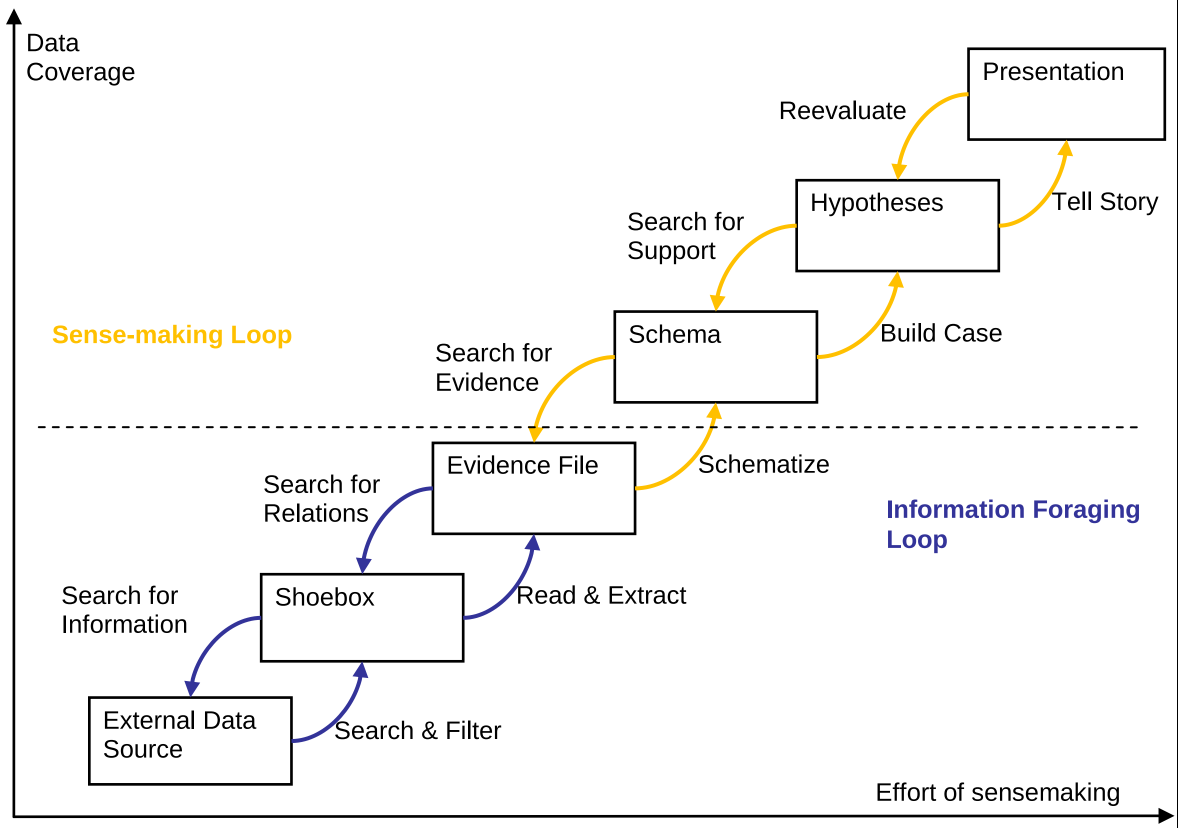


Figure 2-4: Notional Model of Sensemaking Loop for Intelligence Analysis

Recreated from (Pirolli and Card 2005)

The arrows below the dashed line represent the information foraging loop, and the arrows above it represent the sensemaking loop. As the sensemaking process goes on, the products or representations become more structured and the effort required for sensemaking increases.

Box 2-3: Notional Model of Sensemaking Loop for Intelligence Analysis

This model gives a clear illustration of the steps and outputs involved in a complex sensemaking process. However, sensemaking does not always have clear beginning and ending points. The simplified waterfall model runs counter to empirical evidence about several sensemaking tasks, for example, expert decision making (Klein, Moon et al. 2006).

2.1.4 Conducting Research as Sensemaking

The act of conducting research (either quantitative or qualitative) is essentially a sensemaking process. Researchers start with a lack or discontinuance of knowledge, recognize the gaps to be filled (research questions), and conduct research using various methods to bridge the gaps.

Compared to sensemakers in other scenarios, researchers undertake a more systematic approach to identify gaps. Researchers explain specifically what the research attempts to learn or understand by explicitly articulating the research question and sub-questions (Creswell 2003; Maxwell 2005). Once the research questions are recognized, researchers also go through data collection (search) and data analysis and interpretation (sensemaking) processes.

Searching for data is different in research than other sensemaking tasks. The data often does not exist anywhere to be “retrieved”; the researchers need to collect data in experimental (quantitative research) or natural (qualitative research) settings.

Quantitative and qualitative methodologies take different sensemaking approaches. In general, the quantitative approach is logic-driven (deductive, or top-down), starting with a hypothesis, collecting experimental data (often in controlled conditions), and conducting statistical analysis to test if the data supports the hypothesis (Kirk 1995). The qualitative approach, on the other hand, may use a combination of deductive and inductive data analysis (Potter 1996). For example, a researcher may use the Grounded Theory approach, "grounded" in the data and developing increasingly higher level concepts and theoretical models (Denzin and Lincoln 2003). Data-driven approaches are quite common in qualitative analysis, for example, the different coding and display techniques to look for patterns and common themes (Miles and Huberman 1994). Logic-driven analysis techniques, such as looking for rival explanations and deriving a logic model (Yin 2003), may also be used as a cross-examination to ensure the validation of interpretation.

The final product of sensemaking is often a research paper, book, or report, describing the sense made (findings and conclusions). Several intermediate products such as coding schema, case reports, and researcher notes may be produced to assist the sensemaking process.

2.1.5 Organizational Sensemaking

The two types of activities, searching and sensemaking, are also recognized in organizational sensemaking. Choo (2006) frames the organization's adaptability in a dynamic environment into a twofold challenge: sensing and making sense.

Sensing is the noticing of potentially important messages in the environment. The challenge lies in the fact that every part of the environment is interconnected with other parts in complex and unpredictable ways. In this stage, members of the organization attempt to acquire information about events, trends, and relationships in an organization's external environment.

Making sense is the constructing of meaning from what has been sensed about the environment. The challenge for sensemaking is that there are multiple interpretations. Organizational sensemaking is inherently a “fluid, open, disorderly, and social process”.

2.1.6 Comparison and Analysis of the Models

The preceding models all attempt to describe a process whereby knowledge is created. Sensemaking is an individual or collective construction of knowledge. The models are generated from and situated in a different task context and user group, probably describing a slightly different process and with different focus.

Despite the differences, the models have several processes and patterns in common. Table 2-1 compares different sensemaking models and illustrates the common processes.

Table 2-1: Sensemaking Models

			Dervin, 1992, 1998 (general)	Krizan, 1999 (intelligence analysis)	Choo, 2006 (organizational sensemaking)	Pirolli & Card, 2005 (intelligence analysis)	Russell, 1993 (training material design)	Qu and Furnas, 2007 (learning)	Maxwell, 2003; Kirk, 1995 (research)
Task analysis			Gap identification	Task planning	N/A	N/A	N/A	N/A	Research goal and questions
Search	**Exploratory search**		Gap identification and bridging	Data collection	Sensing	Search for & filter information;	N/A	Structural information need	Research Design
	Focused search	Search for & extract relations				N/A	Search for representation	Search for structures	Search for theoretical framework
		Search for data				Search for evidence	N/A	N/A	Data collection
Sensemaking	**Fitting data into structures**			Interpretation and analysis	Making sense	Search for support	Create instances of representations	N/A	Data analysis
	Building structures					Schematize	Search for representation	Explore / identify extract structure	
	Updating knowledge		Gap bridging			Build-case	Consume instantiated schemas	Change, grow, validate structure	Interpretation & conclusion
						Reevaluate hypotheses	Modify representation	N/A	
Preparing task output			N/A	Production and dissemination	N/A	Tell story	N/A	N/A	Writing papers, reports, or books

Several important points are raised in the sensemaking literature:

Sensemaking is comprised of iterative searching (the acquisition of information) and sensemaking (the creation of an understanding) processes.

For the searching or information foraging loop, the evolution of a user's interests depends upon the changing characteristics of the information context. New information gives users new ideas and directions to follow. Users use information from the current situation to decide where to go next (Bates 1989; Ingwersen and Järvelin 2005). Researchers have identified the important role of exploratory search and developed systems to support it (Baldonado and Winograd 1997; Qu and Furnas 2007).

The sensemaking loop, on the other hand, including activities such as skimming, examining details, summarizing, organizing, and identifying patterns, is not as well supported. A key task in sensemaking is to identify patterns of concepts and relationships to build on.

Structure plays an important role in sensemaking: Russell's model (1993) illustrates that the major cost of sensemaking is related to the structural representation, including the cost of building a representation to support required operators in the target task, and the cost of instantiating the representations. The literature seems to suggest that the use of external representations for tasks where the information is too voluminous may be helpful in reducing the costs. A general requirement for sensemaking tools is to capture the structural knowledge created during the sensemaking process, including facts, interim solutions or insights, and their relationships.

To summarize, previous research has identified important processes involved in sensemaking, involving the interrelated activities of the search for information and the creation of an understanding. The importance of structure associated with the processes is recognized: not only do structures influence how people search for information, they are critical to the creation of an understanding. It seems quite clear that sensemakers seek structures in sensemaking tasks and their sensemaking processes are closely related to the structural representation of task situations. However, little is known about the processes and mechanisms that contribute to the creation, modification and update of structures when existing external structures are not available or ready to use. This research aims to examine the creation and update of the structures throughout a sensemaking task.

Tools have been developed to support sensemaking in various ways, mostly to capture intermediate products of sensemaking such as insights (Gersh, Lewis et al. 2006) and analytical thoughts (Lowrance, Harrison et al. 2001), and to provide a workspace of the intermediate representations (Wang and Haake 1997; Hsieh and Shipman 2002; Wright, Schroh et al. 2006). However, there is less support for connecting intermediate products to the conceptual structure that users develop. This dissertation provides the basis for developing sensemaking tools that supports users' structuring a conceptual space using and consisting of various sources, including search results and intermediate structured representations such as concept maps, templates, and outlines. Related tools will be reviewed in section 2.4.

2.2 Task-based Information Seeking

Sensemaking is often embedded in work tasks (as opposed to information tasks) (Ingwersen and Järvelin 2005). Information tasks include both search tasks and sensemaking tasks. Among several task characteristics recognized in the review by Kim and Soergel (2005), the tasks that require at least some degree of sensemaking often involve:

- New situations or problems
- Complex, less structured situations or problems
- A new domain
- An unclear information need

Most sensemaking research involved some type of work task(s). The representations constructed during sensemaking process need to fit the task, or they must be updated (Russell, Stefik et al. 1993). Information tasks and work tasks are compounded. For example, learning (an information task) and decision-making (a work task) are the most studied sensemaking tasks. Baldonado and Winograd (1997) studied a task that requires users to learn about Java to decide whether a system interface can be implemented using Java. Other examples include instructional material design (Russell, Stefik et al. 1993), intelligence analysis (Pirolli and Card 2005), and decision making (Wright, Schroh et al. 2006).

Researchers in information studies examined tasks with various complexities, such as routine information processing tasks, normal information processing task, and decision task (Vakkari and Hakala 2000; Bystrom 2002). Findings suggest that different types of information (task information, domain information, and problem-solving information) were sought for different types of tasks and/or at different stages of the task, for example, background information is sought at the beginning (pre-focus) stage.

Sensemaking has different levels - one could make sense of situations by understanding the existing information (or sense made by others), or one could make new sense from the existing information and create one's own interpretation. The creation of structures may also be related to the nature, complexity, and stages of the task.

Kuhlthau (1993) suggested that information that is relevant in general terms is used at the beginning of the task performance, whereas information that is more specific, more pertinent to a chosen focus, is used at the end of the task. Research in topical relevance (Huang and Soergel 2006) reveals different ways in which a piece of information may be useful to a task. For example, a piece of information may be useful to the sensemaking task because it provides background information about the task or topic; when the task enters the focus stage, a piece of information may be useful when it is directly relevant to the focus of the task. Other relationships include indirect relevance and comparison.

Results from task-based information seeking research suggest that during the task process users' sensemaking process may have different focus, which may require different types of information being processed, different supports for search and information organization, and so on.

To summarize, sensemaking needs to be investigated in relation with tasks. Sensemaking activities may differ in different stages of a task and need different types of support. Sensemaking tools should provide information organization mechanisms that are flexible enough to support different stages of tasks.

This thesis aims to build on these findings to investigate how the different relationships between a piece of information and a task may be used in a sensemaking tool to help users throughout the different stages of sensemaking, focusing especially on how to organize different sources and formats of information based on these relationships in the visual workspace for creating a conceptual structure.

2.3 Cognitive Processes and Knowledge Structures

Sensemaking is a cognitive task which involves the changing of the sensemaker's conceptual model as new information is acquired. This section reviews theories in cognitive psychology that deal with the structure of conceptual models and the processes of conceptual changes.

2.3.1 Types of Cognitive/Conceptual Change

Several types of cognitive/conceptual changes may happen to the mental representation of knowledge. Piaget (1978) identified two types of cognitive/conceptual changes in knowledge acquisition:

- Assimilation: the addition of information to existing knowledge structures
- Accommodation: the modification or change of existing knowledge structures

Following Piaget, the schema theorists proposed that knowledge is structured in the form of schemas, and distinguished three ways in which existing schemas can be modified by new experience or information (Rumelhart and Norman 1981):

- *Accretion:* the gradual addition of factual information within existing schemas; the conceptual schemas do not change in accretion. Accretion may takes place in the form of adding new knowledge when prior knowledge is completely missing, or in the form of gap-filling when prior knowledge is incomplete (Chi 2007).
- *Tuning*: the evolutionary conceptual change in the schemas for organizing and interpreting information. These changes may involve "generalizing or constraining the extent of a schema's applicability, determining its default values, or otherwise improving the accuracy of the schema" to best fit the data.
- *Restructuring*: conceptual changes that involve the radical change of existing structures or creation of new structures. Such radical changes often take place when prior knowledge conflicts with new information. New structures are constructed either to reinterpret old information or to account for new information.

Furthermore, Vosniadou and Brewer (1987) distinguished between weak revision (the modification of existing knowledge structures) and radical restructuring (the reorganization and creation of new knowledge structures).

Researchers in artificial intelligence (Sowa 2006) argue that the structural representations of learning systems, either natural or artificial, can be modified in three ways:

- Rote memory: the simplest form of learning that converts the new information to a network and adds it without any further changes to the current network.
- Changing of weights: weights associated with the concepts and relationships (for example, probabilities in a reasoning network) may be modified with updated information. This is a particular form of tuning.
- Restructuring: the most complex form of learning that makes fundamental changes to the structure of the network.

The following table shows a comparison of conceptual changes recognized in the literature:

Table 2-2: Types of Conceptual Changes

Piaget, Brown, & Thampy, 1978	**Rumelhart & Norman, 1981**	**Vosniadou & Brewer, 1987**	**Sowa, 2006**	**Chi, 2007**
Assimilation	Accretion		Rote memory	Adding new knowledge
				Gap filling
Accommodation	Tuning	Weak-revision	Changing weights	Conceptual change
	Restructuring	Radical restructuring	Restructuring	

The above three types of changes in one's mental model of a situation may take place as one acquires and makes sense of new information. The sensemaker updates his or her internal knowledge representations so that the new information can be incorporated into his or her existing knowledge.

2.3.2 Cognitive Mechanisms

Several cognitive mechanisms are involved in the processes that result in the accretion, tuning, and restructuring of knowledge. Researchers in the areas of reasoning (Toulmin, Rieke et al. 1979; Arthur 1994; Johnson-Laird 1999), reading comprehension (Kavale 1980), and learning (Vosniadou and Brewer 1987), reported several mechanisms that are important to the understanding of information and creation of knowledge. The mechanisms may fall into two broad categories: *inductive (data-driven, bottom-up) mechanisms* and *structure-driven (logic-driven, top-down) mechanisms*. While in general, the mechanisms tend to belong to one category or the other, the distinction between inductive and logic-driven mechanism is not absolute. Some mechanisms may be used in both ways and some mechanisms may not belong to either category.

Data-driven (inductive) mechanisms involve recognizing or matching patterns from data, and building on the patterns of similarity and differences to generalize to the abstract structure of knowledge. In complicated problems where little structured knowledge is available, sensemakers look for patterns, and use the patterns to construct temporary internal models or hypotheses or schemas to work with (Arthur 1994). A list of inductive mechanisms includes, in the order of increasing complexity of cognitive processing:

- *Key item extraction*: the identification in text of key words/phrases or their associated words/phrases such as synonyms and antonyms (Kavale 1980)
- *Schema induction*: the discovery of the regularities in the co-occurrence of certain phenomena (Rumelhart and Norman 1981; Vosniadou and Brewer 1987)
- *Generalization*: making claims about groups based on a sufficiently representative sample (Toulmin, Rieke et al. 1979; Chi 1992)

Structure-driven / logic-driven mechanisms involve using knowledge schemas and logic to make arguments or reach conclusions. Some deductive mechanism include, in the order of increasing complexity:

- *Definition*: defining different aspects of a concept, such as purpose, function and use (Kavale 1980) or using existing definitions.
- *Specification*: specifying as conditions or requirements of a problem or task (Vosniadou and Brewer 1987); for example, illustrating a general claim by a specific facts or illustrate a class with a specific example.
- *Elimination*: eliminating concepts that do not meet certain criteria in certain attributes (Kavale 1980); for example, a sensemaker may eliminate some places from the search list since they do not show any sign of instability.
- *Explanation-based mechanisms, or reasoning from cause*: examining the causal connections of two phenomena (Toulmin, Rieke et al. 1979)

- *Inference:* drawing a conclusion or making a logical judgment on the basis of circumstantial evidence and prior conclusions (Johnson-Laird 1999)

Mechanisms spanning both categories: some mechanisms can be used in either bottom-up or top-down manner.

- *Comparison:* the comparison of a concept to other concepts or ideas (Kavale 1980)
    - *Similarity:* the recognition of common features or attributes shared by concepts (Vosniadou and Ortony 1989)
    - *Differentiation or discrimination:* the recognition of different features of concepts (Vosniadou and Brewer 1987; Chi 1992)
- *Analogy and metaphor*: concepts that are alike may share common features or belong to common categories, may exhibit other common characteristics (Toulmin, Rieke et al. 1979; Vosniadou and Ortony 1989)
- *Classification*: relating a concept to a broader conceptual category and grouping of sufficiently similar concepts (Kavale 1980)
- *Semantic fit*: examining the reasonableness with which a concept appears to fit a certain slot as it relates to the meaning of the knowledge structure as a whole (Kavale 1980); for example, facts that do not fit the existing knowledge structure are more likely to result in changes in the conceptual structures.

Other mechanisms do not belong to either the data-driven or logic-driven approach:

- Socratic dialogue: critical dialogues to facilitate the awareness of inconsistencies in the current schema. Recognition of anomalies can serve an important function in initiating schema restructuring (Vosniadou and Brewer 1987)

These mechanisms may be combined in several ways to enable sensemakers to undertake several processes that result in conceptual changes in their conceptual space. Among them, three important processes affect the knowledge creation and update (Sternberg 1986):

- *Selective encoding* - distinguishing relevant from irrelevant information. This includes selectively acquiring new properties or attributes of a concept, or new relationships of a concept to other concepts.
- *Selective comparison* - deciding what mentally stored information is relevant for solving a problem. This is the comparison of new acquired information to the existing mental models.
- *Selective combination* - combining selectively encoded or compared information in working memory. In the combining process, sensemakers try to fit new information into existing mental models. Conceptual changes such as accretion, tuning, and re-structuring may occur during the combination process.

For accretion to happen, sensemakers need to recognize how well the new information fits into an existing conceptual element in the knowledge. For weak revisions, the concepts themselves do not change their basic meaning. What changes is the location of the nodes in the context of the representation. Such change can result from the concepts having acquired more attributes, certain attributes becoming more or less salient, and so forth. Radical changes or abrupt changes, on the other hand, happen when a concept evolves into a new concept to either complement or replace the existing concept in the knowledge structure.

2.3.3 Relationship to Previous Knowledge

When the new information fits into the existing knowledge structure, knowledge acquisition becomes easier. People experience internal conflict if they are trying to make sense of new information that contradicts their previous suppositions (Anderson 1984).

The **Cognitive Reconstruction of Knowledge Model (CRKM)** (Dole and Sinatra 1998) identifies three important characteristics of the learner's previous knowledge that influence the likelihood of change: strength, coherence, and commitment.

- **Strength** refers to the richness of a learner's existing ideas: are the ideas well formed and detailed, or sparse and fragmented? The stronger the idea, the less likely it is that change will occur.
- **Coherence** refers to the conceptual coherence of the individual's existing knowledge. That is, whether the existing conception provides

an explanation of the phenomenon, idea, or event and fits together all the evidence.

- **Commitment**: Individuals can be more or less committed to their existing conception, regardless of the idea's strength, coherence, or both. Individuals' commitment to their ideas can come from sensory experience, social group membership, or cultural background.

Examining these factors of previous knowledge can shed light on the conceptual changes that sensemakers experience in their sensemaking processes. These aspects of a sensemakers' existing knowledge and its relationship to the new information together may explain in part the different types of changes (or failure to change) happening to sensemakers' conceptual model.

2.3.4 Internal vs. External Representations of Knowledge

Sensemaking, as a cognitive task, can be considered as a distributed representational system with internal and external representations as two indispensable parts (Zhang 2000). In this section, I review both internal representations (also referred to as mental models) and external representations of the cognitive structure of knowledge as they relate to sensemaking.

While this dissertation focuses on individual sensemakers, one might conceive of computer programs for sensemaking, which needs representation (often referred as "knowledge representation"); the distinction between internal and external representation may not apply.

2.3.4.1 Internal Representation

Internal cognitive structures are considered to be the instrument for many cognitive tasks. Research shows that people use map-like structures to make sense of information (Hoffman 1992). Extracting, representing, and analyzing mental models held by people are based on the following assumptions, rephrased from Carley and Palmquist (1992):

1. Mental models are internal representations.
2. Mental models can be linguistically represented.
3. Mental models can be represented as networks of concepts.
4. The meaning of a concept is embedded in its relations to other concepts.
5. The social meaning of a concept is not defined in a universal sense but rather through the intersection of individuals' mental models.

Assumptions 1-4 are relevant to this thesis. Assumption 5 is related to collective sensemaking and is not discussed in this thesis.

Most theories posit that mental models consist of objects and their relationships (Rumelhart and Ortony 1977; Carley and Palmquist 1992; Jonassen and Henning 1996). Among them, schema theory (Wertheimer 1938; Rumelhart and Ortony 1977) claims that personal knowledge is stored in schemas that comprise our mental constructs for ideas. A schema is a package of integrated information on a topic. It is a data structure for representing the generic concepts stored in memory. A schema contains as part of its specification the network of interrelations that is believed to be held among the constituents of the concept of interest. Schemas can be at all levels of abstraction.

Each schema people construct represents a mini-framework in which interrelated elements or attributes of information about a topic are organized into a single conceptual unit. These mini-frameworks are organized by the individual into a larger network of interrelated constructs known as a semantic network, concept map, or graph. These networks are composed of nodes (representations of concepts) and relationships.

Along with others, two types of concepts may be particularly important to sensemaking research:

- *Goal-derived concepts* (Medin and Smith 1984) – specialized concepts created when engaging in goal-directed behavior. This type of concept is particularly useful in defining and accomplishing tasks. For example, a sensemaker may have a concept named "to examine later" to hold pieces of information that she or he is going to come back at a later time. Depending on how the sensemaker divides the task or problem domain, there may be several goal-derived concepts that define the sensemaking approach.
- *Event concepts* – several sensemaking tasks involve the interpretation and analysis of a particular type of concept: events. Also called "scripts", event concepts are representation of stereotyped events, such as going to a restaurant. The properties of script-as-concept would include the actions that comprise the event. A specific story based on

a script can be constructed as an instance of the concept. The contents of a script seem to be hierarchically organized (Abbott, Black et al. 1985). At the top level is the general goal, at the intermediate level are "scenes" which denote sets of actions, and at the lowest level are the actions themselves.

2.3.4.2 External Representation

External representations are very useful for several cognitive tasks. For example, they can serve as memory aids to extend working memory, form permanent archives, and allow memory to be shared. For some tasks, external representation may be more crucial; they form an intrinsic part of the cognitive tasks, without which the tasks either cease or completely change in nature (Zhang 2000).

Information in the external representations can be picked up, analyzed, and processed by perceptual systems. For many cognitive tasks, the interaction of conceptual knowledge between internal representations and external representation is crucial. External representations can give people access to knowledge and skills that are unavailable from internal representations. A broad categorization of external representation given by (Zhang 1997) includes:

- The knowledge and structure in the environment
- Physical symbols, objects, or dimensions
- External rules, constraints or relations embedded in physical configurations

The knowledge and structure may be represented in the following forms:

- Graphical / diagrammatic representations: support operators that can recognize features easily and make inference directly. Diagrams, graphs, and pictures are a few typical types of graphical / diagrammatic representation, used in many cognitive tasks such as problem solving, reasoning, and decision making (Zhang 2000). Graphic organizers (Ausubel, Novak et al. 1978) are developed to provide diagrammatic representations for such cognitive tasks.
- Textual forms: can be more nuanced, and complete, as well as more detailed and precise than graphical/diagrammatic representations.

Both graphical representation and text (especially the act of creating them) support reflection which is required for logical, analytic, rational modes of thought (Flower and Hayes 1981).

With the growth of computer-based information systems, computer-generated displays as external representation can help the quality of complex information processing tasks for certain types of tasks. Much prior work on the role of external representations in individual problem solving has used well-structured problems. Further studies need to investigate ill-structured, open-ended problems.

A variety of schemes for the external representation of mental models share the basic network format, for example: concept maps, schemas, semantic frames, semantic networks, and decision networks. The four basic objects, although mentioned in various terms, are (Carley and Palmquist 1992):

- Concepts: the “ideational kernel” of a mental model or internal representation of knowledge

- Relationships: the tie that links two concepts together
- Statements about two concepts and the relationship between them
- Maps: networks formed from concepts, relationships, and statements

External representations may be consumed by the sensemaker and become part of her or his internal structures. Internal representations may be elicited and stored as external representations. In sensemaking research, some researchers tried to capture the internal knowledge in the users' mind with tools, for example, Lowrance and others (2001) described a system capturing analytic thought using structured argumentation, which explicitly represents:

- User knowledge about the facts and assumptions
- Hypotheses that are drawn from the facts
- Evidence supporting and contracting those hypotheses

They suggested that such tools can be effectively applied to problems where regular assessment must be made, based upon evidence from multiple sources, within a complex and uncertain environment (See Section 2.4 on design of related tools).

2.3.5 Learning Theories as Theories for Sensemaking

Learning is more than the collection of inputs and the production of outputs. The mind has the ability to extract, analyze, synthesize, and formulate received information and stimuli in order to produce things that cannot be directly attributed to the input given (Gredler 2004). Much of the learning activity (rote learning and meaningful learning), especially meaningful learning, has a high degree of resemblance with sensemaking in terms of the cognitive processes that the learner/sensemaker undertakes.

Three learning theories are particularly useful as theories for sensemaking:

- Schema Theory (Rumelhart and Ortony 1977; Anderson 1984)
- Assimilation Theory (Ausubel, Novak et al. 1978; Novak 1998)
- Generative Learning Theory (Wittrock 1990; Grabowski 1996)

Sections 2.3.5.1 to 2.3.5.3 review each of these theories. Section 2.3.5.4 reviews the acquisition of structural knowledge.

2.3.5.1 Assimilation Theory

Meaningful learning, according to Assimilation Theory (Ausubel, Novak et al. 1978) (also known as "the Theory of Meaningful Learning") is a process controlled by the learner in which a new piece of information is assimilated to an existing relevant aspect of the learner's knowledge structure. Meaningful learning is similar to sensemaking in the sense that the process involves the updating of a knowledge structure stimulated by new information. Assimilation Theory makes the following claims about learning:

- The development of new meanings is built on prior knowledge, i.e. relevant concepts and relationships.
- The learner's cognitive structure is organized hierarchically, with more general, more inclusive concepts occupying higher levels in the hierarchy and more specific, less inclusive concepts subsumed under the more general concepts.
- When meaningful learning occurs, relationships between concepts become more explicit, more precise, and better integrated with other concepts and relationships.

To apply claims from the Assimilation Theory to sensemaking, new meanings are developed upon prior knowledge about the task situation. When new information is integrated with the prior knowledge, sensemaking occurs with relationships between concepts becoming clearer and better integrated. Different from science education, the cognitive structure for a task may or may not be arranged hierarchically, depending on the nature and complexity of the task and the subject domain.

As pointed out by (Novak 1998), rote learning and meaningful learning are on opposite ends of a continuum; thus a lot of sensemaking activities may fall in between the two ends of the continuum. Some direct fitting of facts into existing knowledge structure without understanding its relationships may be similar to rote learning.

Ausubel (1978) proposed the idea of an “advance organizer”, which helps learners bridge the gap between knowledge they already process, and new knowledge to be learned. Concept maps can be used together with other representations for capturing and achieving expert knowledge. They can also serve as a platform for sensemaking. For example, explicit structures of the task or problem domain may be visually organized in which sensemakers can put in facts or data.

2.3.5.2 Schema Theory

Schema Theory states that knowledge is stored in human memory as schemas (Rumelhart and Ortony 1977; Rumelhart and Norman 1981; Anderson 1984) with interconnected concepts and relationships, organized in a meaningful way. Main principles of this theory include:

- Schemas change as new information is acquired;

- Prior knowledge is necessary for new knowledge, but in order for learning to happen, learners need to actively build new schemas and revise them in light of new information. Acquiring general knowledge and generic concepts at the beginning stage of learning is important.
- Each individual's schema is unique and dependent on that individual's experiences and cognitive processes.
- Learners feel internal conflict if they are trying to assimilate schemas which contradict their previous suppositions.

In Assimilation Theory, knowledge is postulated to be a hierarchical organization where the learner more or less attaches new knowledge to the existing hierarchy. Knowledge in Schema Theory, however, is not necessarily stored hierarchically but rather represented as networks of propositions that are actively constructed by the learner.

The Schema Theory of Learning emphasizes the relationships between different components of the schemas. In order for learners to learn and remember certain knowledge, it has to be organized mentally as an interconnected network so that the activation of one concept can activate other related concepts.

2.3.5.3 Generative Learning Theory

Generative Learning Theory (Wittrock 1990; Grabowski 1996) suggests that the learner is actively engaged in the learning process, working to construct meaningful understanding of information found in the environment. The model of generative learning emphasizes the importance two types of meaningful relations:

- Between information and experience

- Among the parts of information

According to Generative Learning Theory, comprehension occurs by formulating connections, rather than simply "placing" information into memory or "transforming" information in memory. The key is that the new understanding of the information is created by the learner, rather than the learner modifying external information.

Two types of generative activities were identified according to the Generative learning theory (Grabowski 1996):

- Learning activities that generate or create organizational relationships between different elements in the environment, including creating graphs and tables, creating titles and headings within texts, and forming objectives, summaries, and main ideas
- Generating integrated relationships between the external representation or stimuli and the memory components (internal representation), including demonstrations, metaphors, analogies, examples, pictures, applications, interpretations, paraphrases, and inferences

The two types of generative activities, translated into sensemaking, represent the activity of extracting concepts and relationships from the new information received from the environment, and the activity of positioning the extracted structure into the existing knowledge.

2.3.5.4 Structural Knowledge Acquisition

One of the central concepts related to learning and sensemaking is structural knowledge. Structural knowledge plays an important role in the creation of understanding or modification of existing knowledge. Research found that learners who were given the task of creating a semantic network performed significantly better on other tasks such as relationship judgments (Jonassen and Wang 1993).

Structural knowledge supports higher order thinking in the form of analogical reasoning. The construction of personally relevant knowledge structures is the key of sensemaking. Jonassen and Wang (1993) found that the best task performance on analogy subscale was achieved by learners who work with visual support of a graphical browser and focus on structural relationships. Structural knowledge acquisition improved significantly by focusing the learner's attention on structural aspects of the information in the system. Visual tools help the acquisition of structures.

To summarize, learning theories provide insights into the type of information and relationships that are most useful for knowledge acquisition, which underlies learning and sensemaking. Two types of relationships: the relationships embedded in the new information and the relationships between the new information and prior knowledge, are essential. The acquisition of structural knowledge helps higher level learning and analogy. Information extraction can assist users with the extraction of relationships in the new information. Visualization tools such as concept mapping can help users with structural knowledge acquisition and fitting new information into existing structures. The next section reviews design of related tools.

2.4 Design of Related Tools

This section reviews the design of related tools that may be used for sensemaking tasks. Chapter 7 Implications for Design discusses the design implications for tools that may be used for assisting users with creating, instantiating, representing, and organizing the conceptual space of knowledge, as well as representing, depicting, and recording insights throughout the whole sensemaking process. Tools for supporting exploratory search, which also is an important area for system design, are not discussed.

This section is organized as follows:

Section 2.4.1 reviews the information extraction (IE) technique that may assist users with creation and instantiation of structural elements.

Section 2.4.2 reviews tools for representing knowledge structures, including concept maps, templates, outlines and textual formats and the comparison of them.

Section 2.4.3 reviews tools that help users with their analysis activities such as recording insights, generating and testing hypothesis, and making arguments.

Section 2.4.4 discusses task-specific and task-independent tools.

2.4.1 Information Extraction

Information extraction (IE) is the technique of automatically extracting structured information from text (Grishman 1997), for example, filling in templates in a pre-defined domain. Stimulated in part by the Message Understanding Conference (MUC) and Automatic Content Extraction (ACE) project (Doddington, Mitchell et al. 2004), numerous researchers have contributed to the area and developed toolkits for various sub tasks. This section reviews several extraction techniques that may be useful to assist sensemaking.

2.4.1.1 Named Entity Recognition and Mention Detection

Named Entity (NE) Recognition (NER) is to recognize and classify special categories of terms, such as names of people or organizations, entity names, temporal expressions, and numerical expressions.

Current techniques of NER are able to reach an F-score (the harmonic mean of recall and precision) of 90% or higher (Mikheev, Moens et al. 1999; Zhou and Su 2001; Florian, Ittycheriah et al. 2003; Florian, Hassan et al. 2004). Machine learning techniques, such as Hidden Markov Models (HMM) and Maximum Entropy Models (MaxEnt) are widely used, alone or in combination, in most NER systems to classify unknown terms to the pre-defined categories. Lexical features, part-of-speech tags, orthographic information, affixes and chunk information were also incorporated in most systems (Tjong Kim Sang and De Meulder 2003). NER systems often use extensive “gazetteers” – lists of known names of people, organizations, locations and other named entities. Performances without gazetteer range from 80-90% recall and precision for people and organizations to40-60% recall and precision for locations (Mikheev, Moens et al. 1999).

Recent research (Doddington, Mitchell et al. 2004; Florian, Jing et al. 2006) on entity detection and tracking extends the goal to not only identify named entities, but also nominal or pronominal references, referred as mentions in the Automatic Content Extraction (ACE) project. This is a more challenging task. All mentions of an entity, regardless of its form, are to be detected and collected into classes. Mention detection performances can reach an F-score of about 80% (Florian, Jing et al. 2006).

Since entity recognition extracts major entities in a document which comprise the main content of the document, it may be useful for sensemaking in which sensemakers manually extract the main entities and concepts, and relationships among them.

2.4.1.2 Relationship Extraction

Relationship extraction aims to extract semantic relationships between pairs of entities from text (Doddington, Mitchell et al. 2004). ACE 2004 includes 7 relation types and 23 subtypes (Kambhatla 2006), shown in Table 2-3:

Table 2-3: Relationship Types in ACE 2004

Type	Subtype
ART (agent artifact)	user-or-owner inventor/manufacturer other
EMP-ORG	employ-executive employ-staff employ-undetermined member-of-group partner subsidiary other
GPE-AFF (Geographical Political Entity Affiliation)	citizen-or-resident based-in other
DISCOURSE	N/A
PHYSICAL	located near part-whole
PER-SOC (personal/social)	business family other
OTHER-AFF (PER/ORG affiliation)	ethic ideology other

Recent report on the relationship extraction performances against human-annotated data reports an F-score of about 65% for the above relationship types (Kambhatla 2006).

Extracted relationships, with mentions of entities, if extracted correctly, may be helpful to sensemakers by saving the manual effort of the sensemaker involved in establishing relationships among entities, especially when the volume of text to be processed is huge. However, incorrectly extracted relationships may be distracting or deceiving to the sensemaker and bring noise to the sensemaking task. It is an open question as to how sensemakers use extracted information in their sensemaking task.

2.4.1.3 Template Filling

Template filling is a form of relationship extraction to extract from unstructured or semi-structured text into a pre-defined template. For example, a job template may include slots such as job title, salary offered, company, recruiter, location, required skills and expertise (Califf and Mooney 1999; Ciravegna 2001). Template filling is very useful in this example to create a job database from massive job ads.

Template filling performances vary from frame to frame, and vary from slot to slot within one frame. Some frames and slots are easier than others. In average, the precision performances range from lower 60% to upper 90% (Ciravegna 2001).

Template filling may be useful to sensemaking tasks when the sensemaker works around a central entity or event and needs attribute information about that entity or event. However, template filling does not work on any entity or event.

Templates need to be pre-defined, and the technique works better for entity types which are well-defined and stable over time.

To summarize, automated extraction of entities and relationships may be helpful to users. Sensemaking tasks involve the creation and instantiation of structures, and information extraction techniques can support users with "instantiated structure elements" i.e., entities and relationships along with the texts from which they are extracted. For example, the system may suggest preliminary formal statements for users to examine and filter, saving users the time reading retrieved documents and extracting relationships manually. How to organize and integrate extracted results into the emerging conceptual structure of users remains a question to be investigated.

2.4.2 Representation and Visualization Tools

Information visualization helps the display of structures. Visualization reveals broader patterns within information sets and helps people recognize characteristics of the information set as a whole. In addition to lowering cognitive load by extending working memory, it helps to reveal spatial relationships. Visualization is a broad field; this dissertation focuses on visualization of internal structures /mental structures.

Most information visualization research (Gaines and Shaw 1995; Chi and Card 1999) has been focused on characterizing the information set or the collection. For example, visual spreadsheets are used to analyze the patterns of use of Webs (Chi and Card 1999). Results show that visual spreadsheets using tree-like structures aid users in understanding the relationships between the pieces of information by showing the intermediate results of each operation.

However, in information visualization research, there is less emphasis on visualizing the conceptual structures that users make through sensemaking task. External representations of a user's mental structure are very useful to cognitive tasks as memory aids to extend working memory, form permanent archives, and allow memory to be shared (Zhang 1997; Zhang 2000). Researchers in areas such as education have used concept maps (Novak 1998) as tools to facilitate meaningful learning process, which is similar to sensemaking in many aspects.

Intermediate representations are very important for tasks such as problem solving and sensemaking (Hsieh and Shipman 2002), since they represent the conceptual structures of the users' cognitive process. Structured information is often only a partial abstraction of the information it represents. Formalized data loses certain aspects of the information it intends to represent. An important figure of a visual workspace should combine direct manipulation of visualization and editing.

2.4.2.1 Concept Maps, Graphs, or Networks

Map-based representations are widely used in several disciplines such as artificial intelligence, education, and management. The cognitive mapping techniques are commonly used to reveal cognitive structures, i.e., knowledge systems that individuals or groups used to interpret the problem domain and take action.

Several forms of map-based representation have been created. For example, the following forms of map or network-based representation are widely accepted by the researchers (Lehmann 1992; Sowa 2006):

- **Definitional networks**, also called abstraction hierarchies, emphasize the subtype or is-a relation between a concept type and a newly defined sub-type. They support the rule of inheritance for copying properties defined for a supertype to all of its subtypes.
- **Assertional networks**, sometimes called conceptual graphs, are designed to visualize assert propositions. Some assertional networks have been proposed as models of the conceptual structures underlying natural language semantics. It uses first order logic to express propositions.
- **Implicational networks** use implication as the primary relation for connecting nodes. They may be used to represent patterns of beliefs, causality, or inferences, also called belief networks, causal networks, Bayesian networks, or truth-maintenance systems. It is a special case of assertional Network in which the primary relation is implication.
- **Learning networks** build or extend their representations by acquiring knowledge from examples. The new knowledge may change the old network by adding and deleting nodes and arcs or by modifying numerical values, called weights, associated with the nodes and arcs.

Despite the difference in focus and implementation, all types of map-based and graphic representations help people to determine their relationships to their environments. Maps provide a frame of reference for what is known and believed (Fiol and Huff 1992).

Several software packages belong to the category of representational tool, known as concept mapping software (also referred to as idea mapping or mind mapping). Concept mapping software allows users to externalize their internal representation of a topic, task, or problem. The existing tools by and large provide similar functions with some variation in input and output format, representation of nodes and arcs, and abilities for collaboration. The tools generally allow users to construct, manipulate, and sometimes share their knowledge models represented as concept maps. The tools may be downloaded and installed on a personal computer, or may be accessible through Internet browsers. Table 2-4 summarizes some representative concept mapping toolkits.

Table 2-4: Concept Mapping Tools

Category	**Function**	**Co-mapping**	**CMap**	**Mind 42**	**Free Mind**	**IMind Map**
Basic operations	Add / modify and delete concept (or idea) nodes	√	√	√	√	√
	Add / modify and remove hierarchical links	√	√	√	√	√
	Add/ modify and remove other links		√			
	Add hyperlinks from nodes and links to other resources		√	√	√	√
	Merge nodes		√			
	Annotate map components		√		√	√
Map operation	Create, delete maps	√	√	√	√	√
	Save maps	√	√	√	√	√
	Comparison of two maps		√			
Formatting	Formatting font and line sizes and colors		√	√	√	√
	Spell check		√			√

Category	**Function**	**Co-mapping**	**CMap**	**Mind 42**	**Free Mind**	**IMind Map**
	Include pictures (can also be done indirectly by hyper-linking) and multi-media components		√	√		√
Search	Integrated search			√		
	Search text in maps		√		√	
Collaboration	Multiple users	√	√	√		
	Permission control on map components	√	√			
Presentation and output preparation	Importing from/ Exporting to other formats		√		√	√
	Print preview, presentation builder	√	√			√
Enhanced	Zoom in and out	√		√		√
	drag and drop, expand and collapse	√		√	√	√
	Automatic textual outline					√

Researchers in several areas used concept maps to support many tasks. For example, concept maps are used in science education to promote students' understanding of scientific concepts (Novak, 1998, 2006; Cañas et al., 2005). Concept maps are used to ascertain what learners know at any point in their educational experience. Various tools were built to assist meaningful leaning by individuals as well as collaborative learning. Concept maps are used as research and evaluation tool for science education (Markham 1994). Concept maps are also used as decision aids for management (Fiol and Huff 1992). Important factors such as causal references, strategic dimensions, and organizational structures in management may be represented with cognitive maps to assist decision.

More research is needed to investigate how representational aids may help users to make explicit their internal conceptual structures and connect the structure to various sources and formats of information or interim representations available in the users' workspace.

2.4.2.2 Semantic Frames, Templates

Semantic frames and templates are representation structures for conceptual or world knowledge (Fikes 1985; Lönneker 2003). Semantic frames and templates can hold a large amount of knowledge about a given concept or entity, and can be used to support a knowledge system's reasoning ability. Individual frames or templates are instances of pre-defined entity or event types. Often a frame- or template-based representation includes two major aspects (Fikes 1985):

- Taxonomy description: describes entities and classes of entities in an application domain. Each entity or class is represented by a frame. Frames can have sets of attribute descriptions called slots.
- Attribute description: describes entity attributes (slots). Slots can have multiple values and a set of properties (facets). Several frame-based systems have built-in facets for representing constraints on number of possible values an attribute can have and for indicating the classes to which each value must belong.

A frame system is like a directed graph with labeled vertices and arcs (Lehmann 1992). Compared to network-based representation, frames provide a more detailed view of the data objects representing schemas.

2.4.2.3 Outlines and Text Representation

Outlines provide a simple way for organizing information that could be easily transformed into a textual form of report, which is in many cases the final product of sensemaking. Threaded discussion in many cases has similar representational features as outlines.

The final product of learning or sensemaking takes the form of text representation. Text representation can be more precise, nuanced and detailed than graphical representations. Moreover, the act of producing text representation--writing, including outlining and editing--enables reflective thoughts that make possible the logical, analytic, and rational modes of thinking (Flower and Hayes 1981).

2.4.2.4 Comparison and Co-existence of Representations

Each form of representation manifests a particular representational bias, expressing certain aspects of one's knowledge better than others. Representational bias manifests in two major ways (Suthers 1999; Suthers 2003):

- Constraints: limits on logical expressiveness, and in the sequence in which knowledge units can be expressed;
- Salience: how the representation facilitates processing of certain knowledge units, possibly at the expense of others.

Most relationships may be represented in multiple ways (Suthers 1999). For example, evidential support may be represented as threaded discussion, containment, or graph, as shown in the following figure:

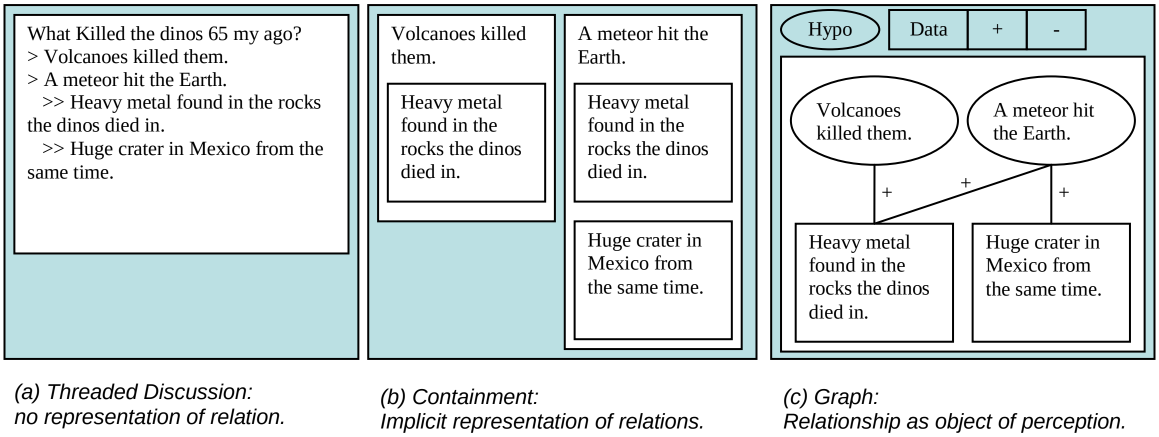
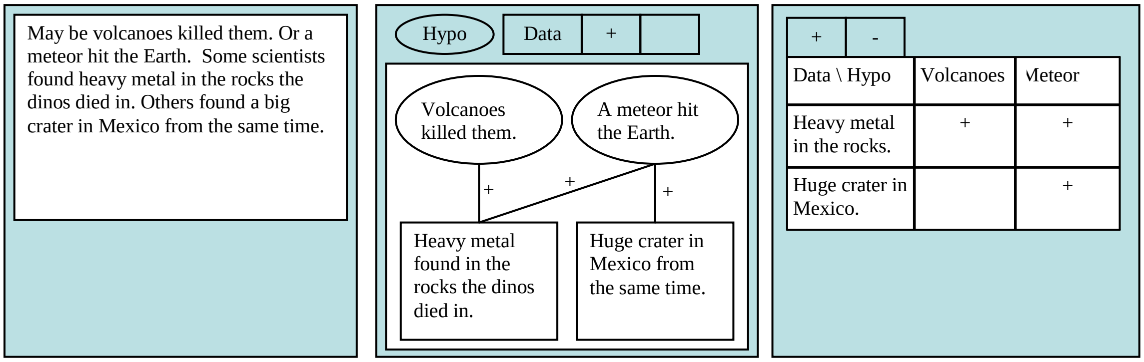


Figure 2-5: Threaded Discussion, Containment, and Graph Representations

The middle representation uses an implicit device, containment to represent evidential support, while the right notation uses an explicit device, an arc. Alternative hypotheses were proposed as to which alternative representation of the relationship will receive more elaboration.

Representation structures that allow unfilled "fields" can make missing knowledge units as salient as those that are present. If the knowledge structure provides structures with predetermined fields to be filled with knowledge units, it may be able to guide search to fill the knowledge gap. Figure 2-6 shows the comparison of three different representations as to their abilities to identify missing knowledge.



(a) Text: No relation is saliently missing.

(b) Graph: Partial salience of missing relations.

(c) Matrix: Salience of all missing relations.

Figure 2-6: Text, Graph and Matrix Representations

The above figure shows three representations that differ in salience of missing evidential relationships (Suthers 1999). Text format often does not help users to identify the missing knowledge. In graph representation, the lack of connectivity shows missing links, but if multiple relationships exist, they are likely to be missed if one relationship is shown connecting the knowledge units. The matrix representation is able to make salient all missing relationships.

Several operational hypotheses were proposed and tested empirically in the domain of computer supported collaborative learning (Suthers 2003), to discover how these notations compare to each other. Table 2-5 illustrates the activities that different representation formats support in a decreasing order.

Table 2-5: Comparison of Representations for Activity Support

Concept Use	**Elaboration on relations**	**Search for missing relationship**
Graph	Graph	Matrix
Matrix	Matrix	Graph
Container	Container	Container
Text or threaded discussion	Text or threaded discussion	Text or threaded discussion

In designing a representational tool, Suthers (2003) depicted a notation / artifact distinction. A representation tool is a software implementation of a representation notation that provides a set of primitive elements out of which representational artifacts can be constructed. The available elements in the notation guide and limit the kind of representational artifacts that can be constructed using the tool.

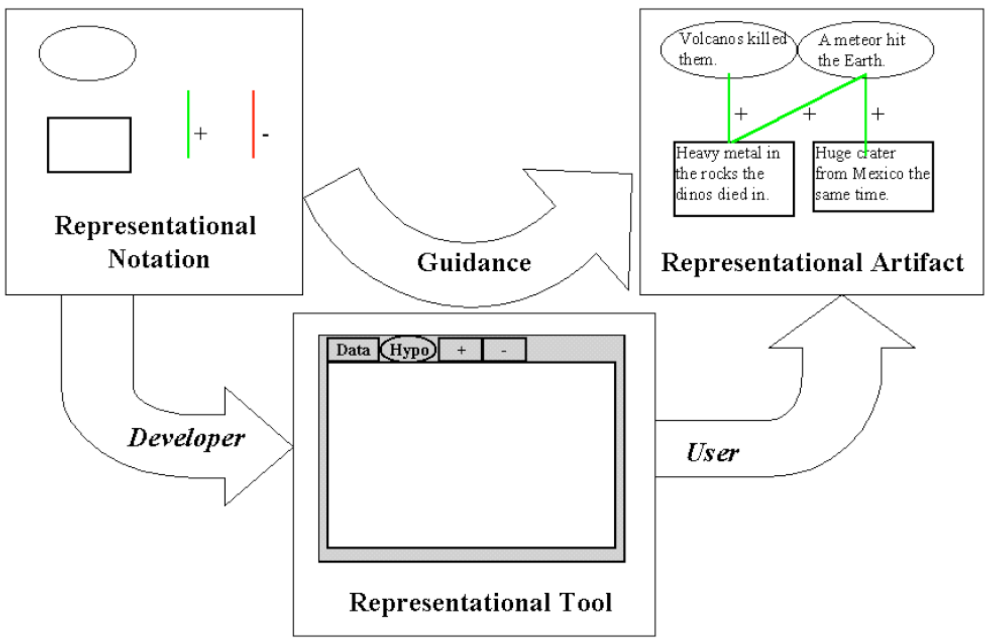


Figure 2-7: Representational Tool (Suthers, 2003)

To summarize, representational tools provide users/learners with the means to articulate emerging knowledge in a persistent medium so that personalized knowledge may be recorded, reused, and shared with others. Different representation formats provide better supports for different activities.

2.4.3 Analytical Tools

Sensemaking tasks can be viewed as finding a path from information → schema → insight → product (Pirolli and Card 2005). From the raw information to the final product of sensemaking, several intermediate processes and products are involved. Most of the intermediate activities involve analysis and synthesis of information. Supporting such activities draws attention from several researchers.

Analytical tools for developing and recording insights, identifying patterns, generating and testing hypotheses may provide support for sensemaking. Some research tries to capture the analytical thoughts in the users' mind with tools. For example, Lowrance and others (2001) described a system capturing analytic thought, which explicitly represented user knowledge about the facts and assumptions, hypotheses that are drawn from the facts, and the evidence supporting and contracting those hypotheses. They suggested that such tools can be effectively applied to problems where regular assessment must be made, based upon evidence from multiple sources, within a complex and uncertain environment. Such tools may be especially useful for problems where formal methods in assessment are difficult to employ, and when decision makers may need to sacrifice structure and rigor.

Evidence-marshalling and synthesis is particularly difficult in some tasks. Wright and others (2006) found that almost equal time was spent doing information retrieval, reference saving and analysis, and organizing/navigating directories and files. To get the overall picture by looking at many pages of text, the analyst relies heavily on memory to connect dots. Users also need support to annotate and organize the domain content according to the sensemaking tasks such as problem-solving or decision-making.

Researchers have also been exploring the support for argumentation support, including discovering evidence, constructing arguments, and testing hypotheses. Cho and Jonassen (2002) developed cognitive tools to scaffold student's seeking of warrants and evidence for supporting claims. The system performs part of the task for the student, and supplants the student's ability to perform some part of the task. Scaffolds are temporary frameworks to support student performance beyond their capabilities, for example, constructing arguments. The argumentation support system defines the following pre-classified types that students could make while solving problems:

- Hypothesis
- Data
- Principles
- Unspecified

They also defined the three types of relationships between the arguments: *for*, *against*, and *and*. Tool bars showing the shapes of argumentation and relationship types are located on top of the scaffold. Results show that scaffolding improves the argumentation quality during individual and collective problem solving.

Similarly, Bell (1997; Bell 2000) developed an argumentation tool for students to support students learning complex concepts. Statements are grouped in containers of hypotheses, and pieces of evidence are associated with each statement. The argumentation tool, SenseMaker, allows students to construct and edit their arguments using a graphical representation. Pieces of evidence from the Web are represented with nodes, described, and grouped using frames.

Other analytical tools include pattern analysis or quantitative analysis of massive data. For example, Romano and others (2000) reported a system that supports analysis of comments of marketing research. The task is to analyze the comments made by Web users of art works, make sense of the comments, and make marketing plan/decisions. The system takes comments from multiple users, and categorizes them for their attitudes/preference for marketing purposes. It also supports collaborative analysis by several analysts.

Despite the effect of designing various analytical tools, the process of knowledge building remains a highly creative activity not mastered by automated means (Hsieh and Shipman 2002). But supports from systems for the creation of knowledge structure at various stages and for activities may help users with such activities.

2.4.4 Task-specific vs. Task-independent Tools

Sensemaking is often embedded in other work tasks such as intelligence analysis and training material design. Some tools have been developed to support task-specific activities, while other tools may be used for various tasks.

Task-specific tools builds task processes into the tool, for example, making argumentations (Cho and Jonassen 2002; Gersh, Lewis et al. 2006). Most task-specific tools were limited for use in certain tasks and are not easily applied to other tasks. Examples of task-specific tools include decision support systems and research and analysis software (such as Atlas TI and NVivo). Often a task-specific tool defines information objects and operations that are specific to certain tasks for manipulation. Users may create representations of their conceptual space populated by instances of these task-specific objects. The activity space instance provides support for creating information spaces consisting of instances of allowed object types, limiting its organization to allowed structures, and potentially offering task specific operations.

Some researchers designed task-independent tools to support certain cognitive activities that are common in several tasks. For example, a user-defined activity space, in which the system does not pre-define objects and relationship for users, can be introduced; users could define object and relationship types related to the specific domain they work in (Wang and Haake 1997). An example of a task-independent tool is the concept mapping tool that can be used for different work tasks such as planning, learning and brainstorming. A challenge for task-independent tools is that the tool needs to capture the common characteristics across different tasks and domains. Another challenge is for the tool to support different formats of representation. Unlike tools designed for a specific task in which only one (or two) format would be sufficient and best fitted, task independent tools need to support different formats of representation, and thus the co-existence and transformation of different degrees of formality, i.e., from very informal and unrestricted representations to very formal representations.

To design a sensemaking tool, it is necessary to separate support for task-specific activities and support for task-independent underlying common activities

2.5 A Comprehensive Iterative Sensemaking Model

This section describes a sensemaking model that illustrates the iterative nature of sensemaking and emphasizes the creation of instantiated structure elements of knowledge. Building on previous sensemaking models, and theories of cognition and learning, the model attempts to show a complete picture of the cognitive processes of sensemaking and provide explanatory power by incorporating the underlying mechanisms and different types of conceptual change.

Several elements are involved (shown in Figure 2-8), including processes, activities, mechanisms and outcomes of sensemaking:

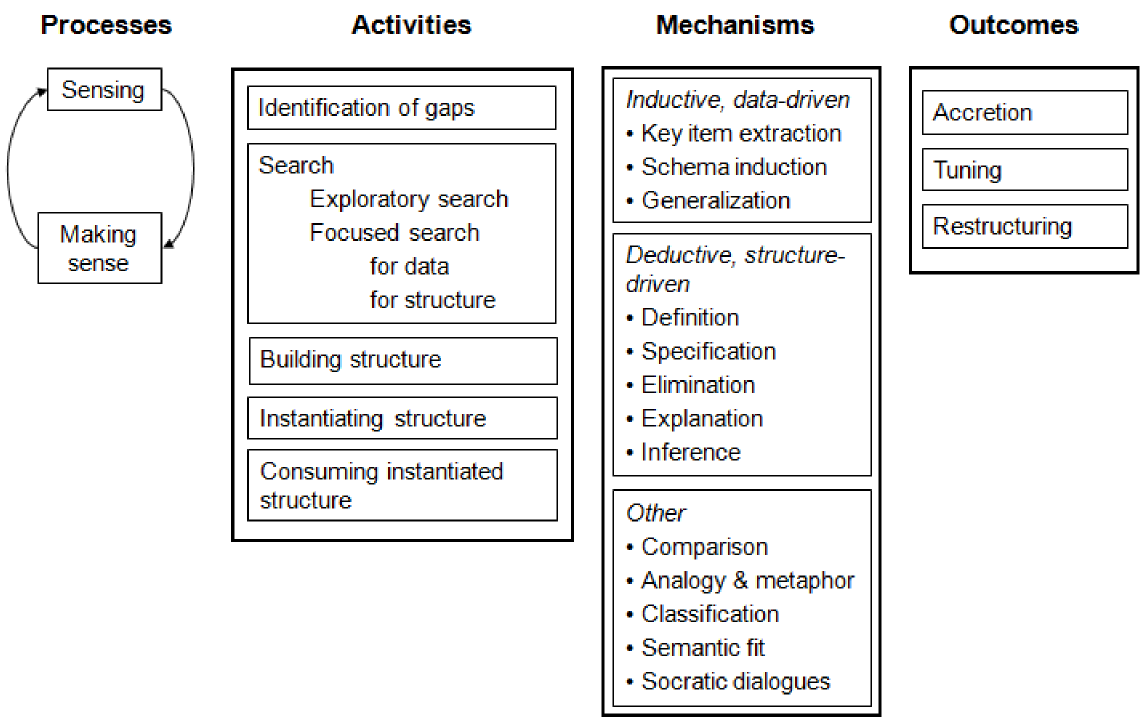


Figure 2-8: Sensemaking Elements

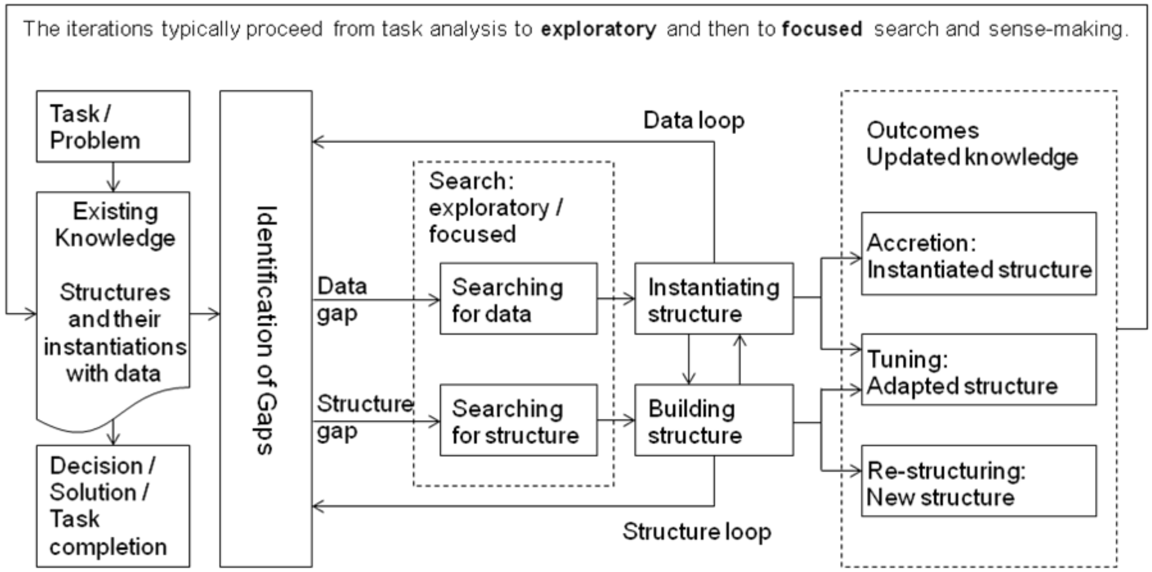
The sensemaking process consists of several iterative loops of searching and sensemaking. Here searching includes both the scanning and monitoring of the environment and the active seeking of information from the environment triggered by problems or tasks at hand. Sensemaking narrowly defined refers to the processes in which users create an understanding or interpretation of what they have sensed consisting of instantiated structures.

Several activities take place within each process. Searching includes pre-focus, exploratory search and focused search for structure and data. Sensemaking includes building of a structural representation, fitting information / data into the representation, and updating of the existing knowledge accordingly.

The activities use several mechanisms, each serving different functions in the structuring of a conceptual space.

The outcome of the sensemaking is an updated conceptual structure which may be updated in three ways: accretion, tuning, and restructuring.

Figure 2-9 shows an iterative sensemaking model that emphasizes the creation of instantiated structure elements and updating of knowledge. Sensemaking involves several processes, which may be executed in many different sequences depending on the level of existing knowledge and the approach of the sensemaker. The model proposed in this paper tries to capture what is typical or, according to the literature reviewed, most frequent.



In each process, the cognitive mechanisms listed below can be used as applicable

Inductive (data-driven, bottom-up)	**Structure-driven (logic-driven, top-down)**	**Both**
Key item extraction	Definition	Comparison
Restatement	Specification	Analogy
Judgment or evaluation	Explanation-based mechanisms	Classification
Summarization	Elimination	*Stereotyping*
Schema induction	Inference	Semantic fit
Generalization		Socratic dialogues

Figure 2-9: An Iterative Sensemaking Model

Sensemaking often starts with the sensemakers' existing knowledge (or the lack of knowledge) of a problem or task situation. The problem or task may be specific or broad. Sensemakers may start with an exploratory search and identify gaps in the existing knowledge, or identify gaps directly by analyzing the problem or planning the task.

Exploratory search is the pre-focus stage of seeking for information. During this process, sensemakers identify a problem, realize they need more information, and learn about what information they need to know through exploring or browsing or broad search. Specific foci have not been established at this stage. Exploratory search may be triggered by an external or assigned task, or it could be continuous monitoring or scanning of the environment. During exploratory search, sensemakers may look for both data and structure, and move through the structure loop and data loop in an embryonic form.

Focused search is a process in which sensemakers search for information about specific aspects of the task situation, having specific questions in mind. The questions represent the gaps identified through problem analysis and/or through exploratory search (also called the pre-focused sensing or exploring phase).

The identification of gaps happens at various stages with different levels of specificity. At the very beginning, the identified gap is a loose notion of lack of knowledge on some topic or problem. As searching and sensemaking continue, more specific gaps may be identified, including data gaps and structure gaps.

If a structure gap is identified, sensemakers may, in varying proportions:

- Search for structures created and described by others and put together a structure from what they found combined with what they already know
- Examine the relationships of various parts of the internal structures in their existing knowledge, look for patterns in the data, and build their own structure or structure modification

If a data gap is identified, the sensemaker conducts focused search looking for the particular pieces of data, and fits the data found into the previously built structure (instantiating representations).

There are two mini-loops involved: the structure loop and the data loop (depending on the focus of a particular iteration of the sensemaking process), both of which are embedded in a larger loop of sensemaking in which knowledge is consistently updated. Sensemakers may take various paths, and the loops may be closely intertwined.

The existing knowledge representation (instantiated structure) may be updated in all three ways: accretion, tuning, and restructuring, referring to both the processes and outcomes of the change.

Instantiating structures may result in accretion (the data fits with the existing structure) or in tuning (the sensemaker makes minor modifications on the structure to let the data fit). Searching for structure may result in tuning (the gradual change in knowledge structure) or in re-structuring (the radical change in knowledge structure).

Some sensemakers may start top down, create structures and then search for data to fit in; others may start bottom up, and any changes in the structure may be accumulated from observing new data. Accumulated accretion may result in tuning, and accumulated tuning may result in restructuring.

Several cognitive mechanisms are involved in the processes; Figure 2-9 gives a preliminary list compiled from the literature. They may be used alone or in combinations. For example, a sensemaker may use the "key item extraction" mechanism to extract key entities, concepts, or relationships as the basic structure elements to build on. She may then use "specification" to specify different aspects or requirements of an extracted concept. Data-driven and structure-driven mechanisms are not the same as the data loop and structure loop shown in Figure 2-9.

The ultimate product of sensemaking is often an updated knowledge representation, which consists of instantiated structures (or schema). The mechanisms described above influence the creation of instantiated structures and the knowledge update. Once the sensemaker incorporates the instantiated structures into his/her existing knowledge, the sensemaking is accomplished.

Chapter 3: Methodology

The purpose of this qualitative user study is to develop, examine, and refine a sensemaking model that is built upon previous sensemaking research, learning theories, cognitive psychology, and task-based information seeking and use. The dissertation aims to understand the complex process of sensemaking, and it focused on the activities, conceptual changes, and cognitive mechanisms used during users' sensemaking process. This chapter describes the methodology and is organized as follows:

Section 3.1 Research design.

Section 3.2 Selection and characteristics of participants and their tasks.

Section 3.3 Data collection methods and procedures.

Section 3.4 Data analysis methods.

Section 3.5 Validity considerations.

Data collection instruments are listed in Appendix B.

3.1 Research Design

To address the central research question "*How* do users make sense of a complex situation (i.e., structure their conceptual spaces about a topic, task, or problem) with the assistance of sensemaking tools?", the research used a collective (Stake 1995), or multiple-case (Yin 2003) design to explore how a sample of 15 participants structured their conceptual space for sensemaking tasks with note-taking and concept-mapping tools. The cases were instrumental as opposed to intrinsic (Stake 1995): they were not the focus of the research but were used to provide insight into a general issue – users' sensemaking process. The multiple-case-study design offers opportunity for replication, which is necessary to develop a theory that can be extended to other cases of individual sensemaking processes (analytical generalization) (Yin 2003; Maxwell 2005).

The case study design is best suited to provide in-depth examination of the complex process of sensemaking, which involves interwoven factors that are best understood using multiple sources of data and "*thick descriptions*" (Denzin and Lincoln 2003).

A small-scale, exploratory pilot study (Denzin and Lincoln 2003) was conducted for two purposes:

1. Instrumental: to test the system environment and to refine data collection instruments
2. Exploratory: to test the ability of the theoretical framework in describing and analyzing the sensemaking process

Three participants from an introductory journalism writing course were recruited by convenience selection (Maxwell 2005). They were working on an assignment of writing an obituary for a living person--in this instance, Maryland congressman Wayne Gilchrest. The task required the participants to gather information about the person from the Internet and write an obituary. Participants worked on the writing assignments using a workstation provided by the researcher. Although the task was more structured than the intended tasks in the main study and the sensemaking involved was limited, the think-aloud data and screen movement recordings still seemed to provide good data for understanding the process. The research design was improved in the following aspects:

1. Changes were made to training materials. The focus of training was changed from techniques (how to use a particular feature or function) to including more workflow, i.e. how a user actually uses the tools for a task. The two-part training (training of tools followed by think-aloud practice) worked well and was carried over to the main study.
2. Instruction for thinking aloud was added (shown in Appendix B.6, Part 2). The researcher experimented by being present for some think-aloud sessions and absent for others, and findings seemed to support what the literature on think-aloud methods have suggested: users were more comfortable talking when they were left alone (Ericsson and Simon 1998; Nielsen, Clemmensen et al. 2002).
3. The interview questions were revised (shown in Appendix B.8) when they seemed to be unclear to the participants.

3.2 Participants and Tasks

3.2.1 Participants

Twenty undergraduate students taking advanced undergraduate courses were recruited from three courses from the Business School and two courses from the Journalism School at the University of Maryland through both *purposeful selection* and *convenience selection* strategies (Miles and Huberman 1994; Creswell 2003; Maxwell 2005). The criterion for selecting students from these courses was that at least one or more course assignments required searching for, analyzing, and synthesizing information, and creating deliverables in the form of a report or a solution to a problem. The types of assignments from these courses included business case analysis and news writing. The assignments are discussed in Section 3.2.2.

Recruitment flyers were distributed to the students in these courses and students signed up for the research on a voluntary basis. (See Appendix B.2) As an incentive to participate in the research, students were paid $15-$20 per hour for their participation. The compensation was provided to the participants for the substantial time and effort they invested. Participation in the study allowed students to learn strategies for organizing information from disparate sources, which may be useful for their future work.

The 20 participants included 6 business students and 14 journalism students. Two participants quit after the training session. Data from 17 task sessions by 15 participants (2 participants participated in 2 sessions) were complete. The following table gives the descriptive statistics of the 15 students who participated fully in the research:

Table 3-1: Descriptive Statistics of Participants

Gender	12 female, 3 male
Age	Range 19-24, average 21
Year in college	2 juniors, 12 seniors, 1 graduate
Major	Journalism (including print and broadcast), government and politics, finance, and marketing (Several students had double majors)
Computer use (hours per day)	2-3 hrs: 5; 3-4 hr 4; 4-6 hr: 4; more than 6 hrs: 2
Search engine and systems	Google, Yahoo! Search, CNN.com, New York Times Online, Washington Post Online, Lexis/Nexis
Confidence in locating information with a Search Engine	Range from 5-7, average 5.6 on a 7-point scale

Most of the students were seniors, meaning that they had more expertise in the subject domain of their major than the average undergraduate. This group of undergraduate students was very computer-savvy. The least hours of computer use per day was 2-3 hours, and went up to more than 6 hours for 2 students. They also had high confidence (averaging 6 points on a 7-point scale, range from 5-7) in the ability to locate specific information using a search engine. Nine students had prior experience with concept mapping as a technique to solve problems, but only one of them used concept-mapping software. Two students used note-taking software.

Details on characteristics of participants are listed in Appendix C.

3.2.2 Sensemaking Task Scenarios

The sensemaking tasks were assignments from the courses mentioned above. The sensemaking tasks were mostly analysis tasks (Hansen and Bystrom 2005; Kim and Soergel 2005), meaning that they often involved understanding of a situation, reasoning / explanation to explain the causal links and potential outcomes, and assessment to form opinions and test hypotheses. Students were asked to choose from a list of assignments that they would like to work on, and bring all related materials they would need with them.

The two major types of tasks were news writing and business case analysis. Boxes 3.1 and 3.2 show an example of each:

Energy and Election News Story

Do your own research using search engines that locate polling results as well as scholarly and journalistic sources, and write a 400-word story about the role of energy, including surrounding factors such as global warming, as an issue in the election. The story can be an overview of the issue, or you can focus the topic to a specific facet of the issue.

Box 3-1: Example Task Scenario: Energy and Election News

Trident Integrated Marketing Communications (IMC) Project

You are developing an integrated marketing communications (IMC) plan for a gum product, Trident, including a TV advertisement and two other advertising and promotion mediums such as print advertising, radio advertising, billboard advertising, direct marketing, web marketing, telemarketing, direct sales, consumer or trade promotions, etc.

Gather current advertisements from your product and its competitors, conduct thorough research in trade and business periodicals on the product, the company, competitors, category users, category trends, and market shares. Develop multiple ideas for the plan based on the research you have conducted.

Your plan should address the problem of the company and the objectives of your proposal, analyze the current marketing and advertising situation, recommend IMC strategy and tactics, and discuss alternatives that you considered but rejected.

Box 3-2: Example Task Scenario: Trident IMC Project (abridged)

In either case, the assignments required students to gather information about the task or problem (either from printed case materials or from online sources), synthesize what they found to create a coherent understanding, and come up with a work product that reflected this understanding.

A list of all tasks is shown in Table 3-2.

Table 3-2: List of User Tasks

Task Title	Brief Description	User Code	Course No.
P&G	Investigating retirement investment plans for P& G employees.	MB1	BMGT440
Boeing 7X7	Analyzing Boeing investment options in its 7X7 models.	MB1	BMGT440
Marketing research	Learning about quantitative marketing research methods.	MB3	BMGT452
"Be Our Guest"	Analyzing business development plan for a small equipment rental company.	MB4	BMGT440
Trident IMC Project	Developing a marketing plan for Trident gum.	MB5	BMGT450
Wireless phone	Developing a marketing plan for a regional wireless phone company.	MB6	BMGT450
Local Election Story	Writing a news story about how a local community thinks about the election.	MJ1	JOUR320
Energy and Election Story	Writing a news story on the role of energy in the 2008 presidential election	MJ3 – MJ15	JOUR471

3.2.3 System Environment

During the training and task sessions, participants worked with a two-screen workstation provided by the researcher. The required software, including standard Web browsers (Internet Explorer and Firefox), MS OneNote, and IHMC CMap, is installed on the workstation.

The general work flow involves the following phases:

1. Searching for information using a standard Internet browser such as Firefox or Internet Explorer;

2. Taking and organizing notes in MS OneNote 2007;
3. Organizing thoughts and notes for the news story or case analysis in CMap;
4. Writing the news story or case report in Word 2007.

Often users went back and forth with the programs using the two-screen workstation provided to them. At any time two programs can be viewed simultaneously. The shortcuts for these programs are listed on the taskbar next to the "Start" menu. Some features of the system environment include:

Taking notes with OneNote. Notes are organized in a "notebook → section → page → (subpage →) note" structure. A note is usually a box of some text with formatting, but a note can also be a picture, or a table. It is much easier to work with Internet Explorer or Firefox on one screen and OneNote on the other screen. Users can highlight a paragraph in the Internet browser window and drag it into OneNote. The highlighted paragraph is put in as a note box and the source (for example, URL of a Web page or file location of a document) is automatically attached to it.

Organizing notes visually. Notes may be arranged and moved freely on a note page. Arrows, boxes, ovals and other shaped can be drawn around the note boxes. Notes can be highlighted and formatted with the standard Microsoft Office formatting options.

Search Notes. Notes can be searched using standard keyword search functions.

Creating and editing a concept map in CMap. Concept maps are organized in folders similar to the Windows directories. In CMap, a map contains two types of elements: nodes and links. A node can be used to represent a concept, an entity, or an event. A link can be used to represent the relationship between nodes. Nodes and links can be formatted with different color, shape, and/or line weights.

Attaching source URLs and/or notes in MS OneNote to a node. URLs can be attached to a node in the concept map by right-clicking on the node and entering a label and the URL. Notes from OneNote can also be attached to a node by getting the hyperlink from OneNote (on the right-clicking menu of a note box or several boxes) and entering it as the URL. An icon appears after a source is attached. When clicking on the icon the system will go to the URL or note in OneNote being attached to this node. If OneNote is not open, it will automatically open the OneNote application and go to the note that is attached.

Participants may also use other computer applications such as MS Word and Excel when needed. The two-screen setup was intended to reduce the number of switching back and forth between applications used. For example, participants often put Web browser on one screen and OneNote on the other so that they were able to search and take notes seamlessly.

3.3 Data Collection

To avoid systematic bias, the research used several data-collection methods: think-aloud protocols, recordings of screen movements, interviews, questionnaires, and document analysis (Potter 1996; Creswell 2003; Yin 2003; Maxwell 2005). The combination of these data collection methods provided a complete picture of the users accomplishing sensemaking tasks with the assistance of computer tools.

Section 3.3.1 describes the data collection methods. It is organized from main to supplemental data.

Section 3.3.2 describes the data collection procedure. It illustrates how the data collection was done chronologically.

3.3.1 Data Collection Methods

The data collection methods involved were:

1. Think-aloud protocols recorded as users work on their tasks. These protocols shed light on users' evolving sensemaking processes, especially on cognitive mechanisms used, and users' internal representation of the knowledge structures.
2. User activity logs automatically recorded by screen capture software (Camtasia). The recordings, along with the think-aloud protocols, captured users' interaction with the system and changes users made in notes, concept maps, and other artifacts from which one may infer changes in the users' conceptual space.
3. Documents produced for the tasks, including notes users took, maps they created, and the final work product which could be a news story or a case report. The documents reflected the external representation of the knowledge structures.
4. Interviews. A post-session interview was used to learn about users' sensemaking strategies and approaches to the task. An exit interview at the end of the semester during which the research was conducted

was used to learn about the long-term impact on users' sensemaking approaches.

5. A user background survey provided information about users' demographics, computer skills, and problem solving skills/approaches.

Multiple sources of data relevant to the same processes were collected and analyzed for triangulation to ensure validity.

3.3.1.1 Think-aloud protocols

This thesis aims to understand users' cognitive processes in sensemaking, which cannot be easily observed at the behavior level. Think-aloud protocol analysis (Ericsson and Simon 1993; Ericsson and Simon 1998) has been widely used in several domains to elicit the cognitive processes responsible for users' behavior, such as cognition (Nisbett and Wilson 1977), education (Chi 2007), instructional system design (Jonassen and Henning 1996; Jonassen 2005), and human-computer interaction (Nielsen, Clemmensen et al. 2002). These authors consider think-aloud protocol a valid source for discovering and verifying human cognitive processes.

The main methodological concern is to collect data about the associated thoughts without altering the structure and course of the naturally occurring thought sequences. Research in cognition (Ericsson and Simon 1998) found that when participants are thinking aloud, their sequences of thoughts are not systematically altered by verbalization. However, when they are asked to explain or describe their thinking to another individual, their performance is often improved.

To ensure that the participants had minimal reactive influences on their thinking, they were explicitly instructed to focus on the task while thinking aloud and merely to verbalize their thoughts, rather than describe or explain them to anyone else. The participants were also given a training task, preparing the participants for the think-aloud exercise and familiarizing them with the system. (See Appendix B.6 Part 2 for the think-aloud instruction and training task.)

Individual differences exist with regards to the ability to perform tasks while expressing one's thoughts. Some may be less verbal than others. Silence may occur and become another challenge (Nielsen, Clemmensen et al. 2002; Sharp, Rogers et al. 2007). Besides training, for participants who were unable to provide verbal reports of their thought process, the researcher prolonged the post-session interviews with less-verbal participants, asking participants to reconstruct several critical incidents and critical threads involved in the sensemaking process, using the interview data as triangulation to ensure validity.

Think aloud protocols were recorded along with screen movements using a program (Camtasia). A stand-alone audio recorder recorded the think-aloud protocols to minimize potential data loss.

3.3.1.2 Recordings of Screen Movements

Screen movements including all search, note-taking, concept mapping, and writing activities were automatically recorded by a program (Camtasia):

- Search engine / system used and queries issued to the system
- Results retrieved from the system
- Documents examined
- Note-taking activities using OneNote

- Addition, modification, and deletion of concepts and relationships in the concept maps
- Writing activities using Word

The software recorded what happened on the computer screen (both screens in one frame) along with the think-aloud audio input from the participants. The recording was exported as a video file (.avi) and played using a regular media player or the Camtasia player.

Participants were informed that their screen movements and think-aloud protocols would be recorded. Running the recording software did not affect the performance of the computer, and was noticeable only through a small icon in the status bar.

3.3.1.3 Interviews

The post-session interview was intended to learn about how participants perceived the changes that happened to their understanding of the topic of the task, how they thought about the tool, and to provide a chance for participants to give input to system design.

It also served as another source to verify the think-aloud protocols generated by the participants during the process. Participants were asked to recall some critical instances of when and how their conceptual model changed.

The post-session interview was semi-structured. The researcher used an interview protocol with the main questions, and followed up on interesting issues that the participants mentioned. Lengths varied from 20 minutes to 40 minutes. See Appendix B.8 for the interview protocol.

The exit interview was intended to learn about the long-term influence of the experience of using sensemaking tools for tasks on the participants' approaches to similar tasks, and to get suggestions from users on designing tools and learning about sensemaking skills and tools. It was arranged at the end of the fall 2008 semester during which the study took place. The exit interview was also semi-structured. Lengths varied from 20 – 40 minutes long. See Appendix B.9 for the exit interview questions.

3.3.1.4 Questionnaires

The User Background Questionnaire was administered as part of the screening process during participant recruitment. The user background questionnaire collected information about users' demographic information, as well as background information that was relevant to the study, such as educational background, computer skills, and problem-solving skills. Refer to Appendix B.4 for details.

3.3.1.5 Documents

Documents including case materials, intermediate products (notes and concept maps), and final products of sensemaking (a new story or a case report), were collected.

Participants were asked to bring all printed materials they needed for a case, and the researcher made copies of printed materials and saved electronic documents when possible. At the end of the session, participants were asked to save their notes, concept maps, and final write-up. They were allowed to keep a copy of the task output for their own purposes if they desired to.

3.3.2 Data Collection Procedures

Participants were asked to attend at least two sessions: a training session and an assignment session. The sessions were conducted in September 2008 and October 2008, varying from 1.5 hours to 3 hours in length. In total, data from 17 assignment sessions was completed.

3.3.2.1 Training Session

Training was done one-on-one to make sure the students following the training instructions step-by-step and their questions answered. At the beginning of the training session, users were informed about the purposes and procedures of the research. They were asked to sign an IRB consent form after the researcher explained the research to them. (Appendix B.3) Users also filled out the user background questionnaire at the training session (Appendix B.4).

In the one-hour training session, participants were given the instructions and a training task to learn how to use the software (OneNote for note-taking and CMap for concept mapping) for the first forty minutes. In the second part of the training (twenty minutes), they continued to work on the training task and practiced thinking aloud. A researcher was present throughout the training session to give instructions, answer questions, and probe the participants occasionally when the subjects were not talking as they did the think-aloud exercise. The training task was designed to be similar to the task that the users were to perform in the assignment session. The business students completed a training task of business analysis and the journalism students of writing a news article. (Appendix B.6)

For each training session, the researcher checked an item on the training objectives (Appendix B.5) whenever a user achieved the training objective as the training moved along, to make sure that the users learned the basic functionalities of the tools.

3.3.2.2 Task Session

The task session was usually scheduled within two to four days after the training session to make sure that the participant had enough time to absorb the training materials and yet had a fresh memory of the training when s/he come to the task session. The business case analysis assignment sessions were scheduled for two hours. The students worked on their cases for about 1 hour and 40 minutes. The news writing assignment sessions were scheduled for 3 hours. The students worked on the assignment for about 2 hours and 40 minutes. Some users did not require all of the scheduled time to finish the assignment. Only one user did not finish the task during the scheduled time.

Participants were asked to bring any printed materials that they might need for their tasks. At the beginning of each task session, the researcher explained to them what they were to do, gave them an instruction on think-aloud and the training instruction in case they needed to look something up, and left the participants to work on their tasks. They were asked to think aloud while performing the tasks.

The business students were asked if they would like to come back for a second assignment session. All students said that they would like to come back for another case session. For scheduling reasons, only 2 out of the 5 students did a second business case session. One continued to work on the assignment and the other worked on another case.

The 20-minute post-session interview was conducted at the end.

3.3.2.3 Exit Interview Session

At the end of the 2008 Fall Semester (during which the data collection was done), an exit interview was conducted to learn about the long term influence on participants' sensemaking approaches and get suggestions for sensemaking tools design and education. Eleven students (4 from business students and 7 from journalism) participated in the exit interview.

3.4 Data Analysis

In this study, a case defined as one sensemaking session, i.e., a user working on a task with the assistance of the system, is the unit of analysis (Yin 2003).

Interviews and think-aloud protocols were transcribed. Researcher notes were inserted to the transcriptions. Transcriptions were coded against an initial coding scheme. The initial coding schema was constantly updated. Emerging patterns and themes were noted and added to the coding scheme. Questionnaire answers and task outputs were analyzed. All types of data collected about each case were put together to generate the individual case report. Multiple sources were used to ensure the validity of the analysis. Conflicting evidence, if any, was reported.

After the individual case analysis was done, the researcher conducted cross-case comparisons by participant and task to discover any common patterns in multiple cases.

3.4.1 Analysis of Think-aloud Protocol and Screen Recordings

3.4.1.1 Transcription

The session recordings (including think-aloud protocols and screen movements) were transcribed according to the Transcription Guideline and Conventions listed in Appendix D. This guideline was developed and updated by two researchers as they moved along with the transcription. The guideline specified the format of the transcript, the prosodic effects of the think-aloud, and the types of user activities.

Description of user activities as recorded by the screen recording software (Camtasia) and the think-aloud protocol were aligned to time spans, which were segmented at natural logical breaks (such as switching of applications and breaks in think-aloud) of the sensemaking sessions. Two researchers transcribed one case independently and compared the boundaries of the sessions. 86% of the boundaries matched. One researcher had finer segments than the other which caused most of the mismatches. The two researchers then continued with the rest of the transcription.

The following table shows part of an example transcript. The “User activity” Column describes the user actions, including activities happening in Web browser, the note-taking application (OneNote), the concept mapping application (CMap), Word, and any other program that were used for the sensemaking task. A list of activities transcribed can be found in Appendix D. The “Think-aloud protocol” column reports the thinking aloud data as the user was involved in the activity shown during the sensemaking activity in the corresponding time span.

Table 3-3: Example Transcript Segment

Time	**User activity description**	**Think-aloud protocol**
…	…	…
00:03:01: :00:03:30	Clicks back to previous Google search results and searches on "energy poll".	I'm going to go back <sp /> and maybe do "energy poll" maybe. <sp /> That's from March. Um, don't know how relevant that is to the election now. Things have changed drastically since then. Might as well check it out though.
00:03:31: :00:04:06	Clicks "pollingreports.com" link. Scrolls up and down page reading poll results. Reads CSS/Opinion Research Corporation Poll re: increase in gasoline price affect on household finances.	<sp /> "Have recent price increases in gasoline pricing produced any financial hardship on your household?" Alright. Most people are saying "yes". <sp /> Oh this September, no, wait, September of '08. That's pretty recent. So yeah, a lot of people are saying it affects them. Which might have to do with the election but not directly
00:04:07: :00:05:03	Scrolls down page, moused over CNN/Opinion Research Corporation Poll re: increased drilling for oil and gas offshore in U.S. waters.	OK. This could definitely have to do with the election here because the republicans are favoring offshore drilling. So, the most recent poll would this one in August. And it looks like a lot people strongly or mildly favor increased drilling for gas in offshore waters. <ironic>Which is interesting.</ironic> Um. I'm going to… Ok. Ooh this is interesting because it's Alaska and Sarah Palin is from Alaska "Do you favor or oppose allowing drilling in the Arctic National Wildlife Refuge in Alaska. A lot of people are favoring that. Um. Ok that's really interesting.
00:05:04: :00:05:13	Selects and copies poll info in browser.	OK, I will put both of these into my notes. Let's see.
…	…	…

3.4.1.2 Coding Scheme

Think-aloud protocols and interviews were coded using an initial coding scheme that emerged from the literature, and emerging themes and patterns were noted and added to the coding schema throughout the analysis process (shown in Boxes 3-3, 3-4, and 3-5).

A Activities

A1 . Search

A1.1 . . Exploratory search

A1.1.1 . . . Search for data

A1.1.2 . . . Search for structure

A1.2 . . Focused search

A1.2.1 . . . Search for data

A1.2.2 . . . Search for structure

A2 . Sensemaking

A2.1 . . Gap identification

A2.1.1 . . . Data gap vs. structure gap

A2.1.1.1 . . . Data gap

A2.1.1.2 . . . . Structural gap

A 2.1.2 . . . Dealing with gaps

A 2.1.2.1. . . . Search for information immediately

A 2.1.2.2. . . . Search for information in the next iteration

A 2.1.2.3. . . . Record the gap and move on

A 2.1.2.4. . . . Ignore the gap

A 2.1.3 . . Ways of gap identification

A 2.1.3.1. . . . By task analysis

A 2.1.3.2. . . . By questioning / Socratic dialogue

A 2.1.3.3. . . . Ad-hoc gap identification

A2.2 . . Building structures

A2.2.1 . . . by adapting others' structure

A2.2.2 . . . by task analysis

A2.2.3 . . . from prior knowledge

A2.2.4 . . . from new information

A2.3 . . Instantiating structures

A2.4 . . Updating knowledge

Box 3-3: Coding Scheme – Activities

B Conceptual changes

B1 . Types of conceptual changes

B1.1 . . Accretion

B1.1.1 . . . Accretion by copied notes

B1.1.2 . . . Accretion by restated notes

B1.1.3 . . . Accretion by summary notes

B1.1.4 . . . Accretion in concept maps

B1.1.5 . . . Accretion to internal structure

B1.2 . . Tuning

B1.2.1 . . . Changing a concept or category's scope

B1.2.2 . . . Changing weight of a concept

B1.2.3 . . . Changing the relationship between two concepts

B1.3 . . Re-structuring

B1.3.1 . . . Creating a note page

B1.3.2 . . . Creating structures inside a note page

B1.3.3 . . . Creating new concepts and relationships in CMap

B1.3.4 . . . Deleting concepts and Relationships

B1.4 . . Changes in representation

B2 . Sensemaking success and failure

B2.1 . . Sensemaking success

B2.2 . . Sensemaking failure

Box 3-4: Coding Scheme – Conceptual Changes

C Cognitive Mechanisms

C1 . Cognitive mechanisms by functionality

C1.1 . . Mechanisms used in processing new information

C1.1.1 . . . Key item extraction

C1.1.2 . . . Restatement

C1.1.3 . . . Judgment or evaluation

C1.1.4 . . . Summarization

C1.1.5 . . . Generalization

C1.1.6 . . . Schema induction

C1.2 . . Mechanisms in examining concepts

C1.2.1 . . . Definition

C1.2.2 . . . Specification

C1.3 . . Mechanisms used in examining relationships

C1.3.1 . . . Comparison

C1.3.2 . . . Analogy and metaphor

C1.3.3 . . . Stereotyping

C1.3.4 . . . Classification

C1.3.5 . . . Explanation-based

C1.3.6 . . . Inference

C1.4 . . Mechanisms used in examining anomalies and inconsistencies

C1.4.1 . . . Elimination

C1.4.2 . . . Semantic fit

C1.4.3 . . . Socratic dialogue

C2 . Data-driven vs. Structure driven

C2.1 . . Data-driven mechanism

C2.2 . . Structure driven mechanism

C2.3 . . Mechanisms that can be used in both ways

C2.4 . . Other

C3 . Cognitive mechanisms by complexity (number of items to deal with at a time)

C3.1 . . One

C3.2 . . Two

C3.3 . . N

Box 3-5: Coding Scheme – Cognitive Mechanisms

3.4.1.5 Plotting Sequence Chart

The processes and knowledge updates were plotted along the timeline during which the session took place. Figure 3-1 shows a segment of an example transcript. The pink bars in the middle illustrate the sequence of search and sensemaking activities. The whole session was divided into several search-sensemaking iterations based on the sequence as numbered in the top row. The knowledge updates and cognitive mechanisms used were plotted as dots in the grid. For example, 4.5-5 minutes along the timeline, an instance of accretion (conceptual changes) was seen when the user was instantiating a structure with data (processes), and the user used summarization (cognitive mechanism) to put data into the structure.

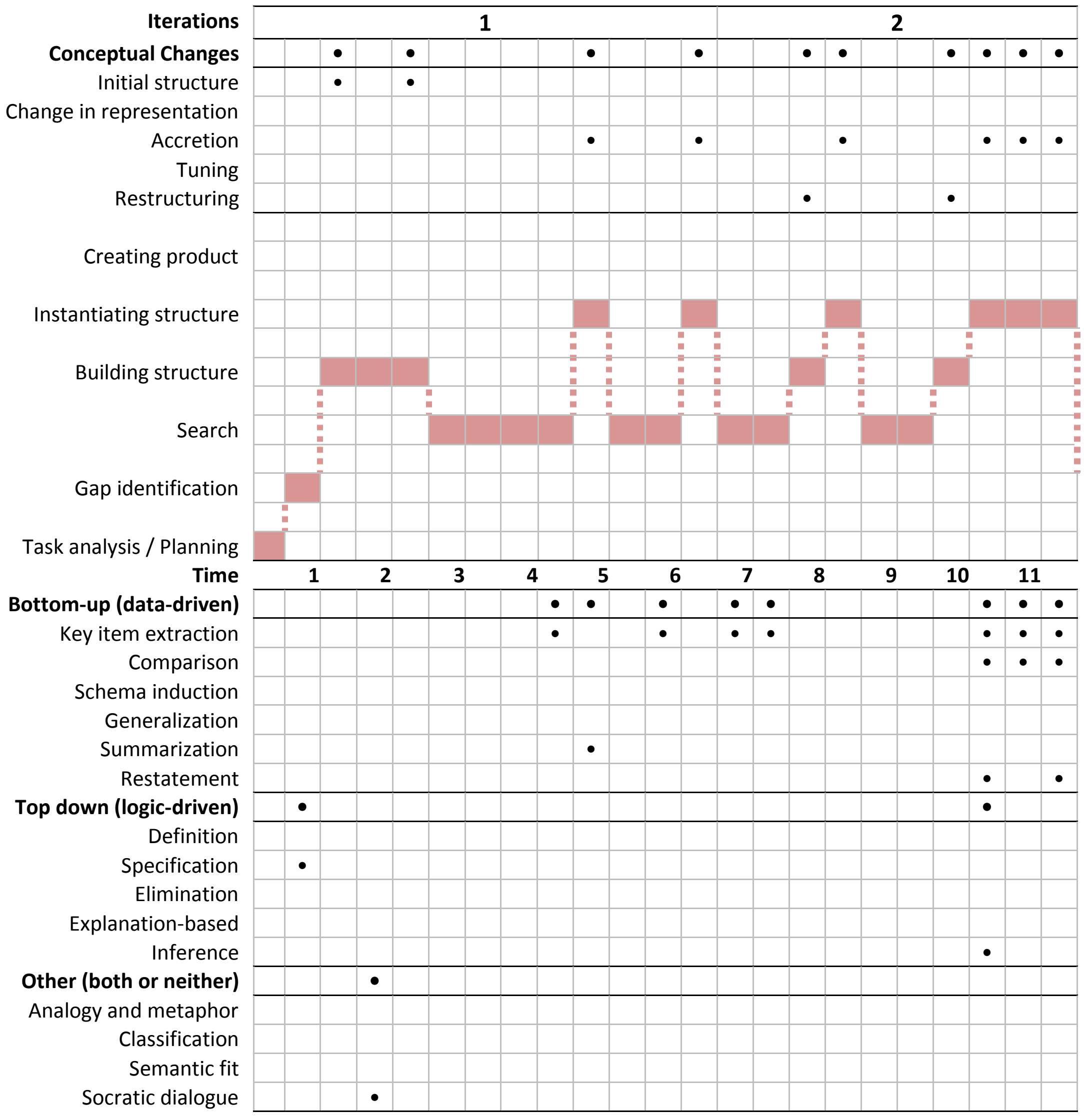


Figure 3-1: Sequence Chart Segment (MB5, Case 1)

Patterns of paths taken were recognized. For example, two patterns of paths emerged from the above example sequence:

1. A full-fledged path from gap identification and search to instantiating structure and building structure:

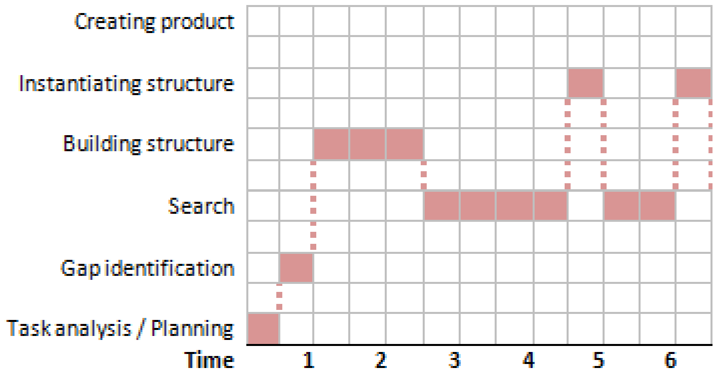


Figure 3-2: Example Path 1

2. A skimpy path of two successive "search + building structure + instantiating structure" sequences:

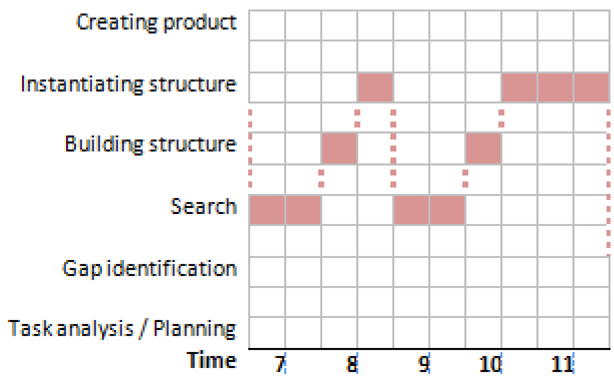


Figure 3-3: Example Path 2

3.4.2 Analysis of Background Questionnaire Data

Descriptive statistics user background questionnaire was computed. It is reported under Section 3.2.1, Participants. Users' demographical, educational, and working characteristics were considered when analyzing the transcripts and task outputs to better understand the processes and consequences. Detail about user characteristics is listed in Appendix C.

3.4.2 Analysis of Interviews

Post-session interviews and exit interviews were used as triangulation to the think-aloud and session recording data, as well as major inputs of user comments on the sensemaking tools used in the study. Interviews were transcribed and coded. The main categories for coding include:

- Tool design
  - Usability
  - Features
- Sensemaking approach and skills
  - For news writing
  - For business analysis
  - For other tasks
- Sensemaking Education
  - Of skills
  - Of tools

Box 3-6: Coding Categories for Interviews

3.5 Validity Considerations

3.5.1 Limitations

Because of the nature of the qualitative user study, one limitation of the research is that the results were transferable, but not generalizable. No claim was made about how representative the participants were to college students nor to the domain workers they were trained to become (journalists and business analysts).

However, these computer savvy users with moderate domain knowledge showed some common patterns in their sensemaking process. These patterns are transferable to other users groups with similar characteristics.

The activities that sensemakers involved in while creating the knowledge structures were influenced and constrained by the software tools that they used in the study. The capabilities and limitations of the software tools may have influenced the process. For example, some users might have taken more or less time to write the news story than they would normally do. For another example, users' concept maps may be represented in different formats given a different concept mapping tool. Because of this, the research did not intend to make any claim about the time allocation for different activities nor about the use of any specific software. Findings on the sensemaking patterns held true regardless of the specific software used.

3.5.2 The Involvement of Other Researchers

Two other researchers were involved in the research. A Master student in the College of Information Studies conducted a few of training sessions, and coded two think-aloud and screen recording transcripts (the main researcher coded all transcripts) to examine inter-coder reliability. A third researcher, who is the chair of the dissertation committee, conducted the exit interviews.

Although multiple interpretations were possible when multiple researchers were involved, both additional researchers were familiar with the theoretical framework of this dissertation, and discussions were conducted regularly so that all researchers developed shared understanding to guide the research activities they were involved in.

3.5.3 Triangulation

Data from multiple sources was collected and analyzed in relation to other data. Together, they provide a complete and detailed picture of the users' sensemaking processes. The think-aloud protocols and screen movements were the main data sources, correlated by the time stamp that the screen capturing software (Camtasia) recorded. In post-session and in exit interviews the researcher asked users to describe the general processes they went through so that the interpretation of the think-aloud and screen movements could be cross-validated. The researcher also asked the participants to explain their maps and notes as another way to verify the interpretation of the task outputs and conceptual changes. User background and characteristics were useful in understanding the user context and for the researcher better interpreting the results.

3.5.4 Inter-coder Reliability

To test the stability of the initial coding scheme, a second coder coded a sample of two think-aloud transcripts. For the first transcript, the two coders coded the transcript according to the initial coding scheme independently and the agreement was 63%. Then the two coders come together to resolve any disagreements in coding; the disagreements occurred mostly on cases where a user used closely related cognitive mechanisms. The discussion helped in defining these codes. Then the two coders proceeded in coding a second article independently and reached an agreement level of 87%. The researcher proceeded with the rest of the coding.

3.5.5 Case Report and Member Checks

A case report that was created for each case after all the analysis about the case was completed to summarize and synthesize the findings. A case report includes (see Appendices D and E for detailed examples):

1. A task / assignment description
2. Description of the "search-sensemaking" iterations that users went through
3. Think-aloud protocol and description of user activity with coding (the examples in the Appendices D and E are re-created from NVivo)
4. Note pages and note structure
5. Concept map(s)

Member checks were conducted to ensure the validity of the research. Since participants who were seniors at the time of the data collection graduated when analysis was done, only two participants were approached for member check (MB5 and MJ3). The participants received a case report summarizing the findings and were asked to provide any feedback they had about the findings. They were asked to rate from 1-5 on how accurate each of the items listed above from the case report reflected their experience in the task session. If an item was rated 4 or under, they were asked to give an explanation of what was missing or incorrect so the researcher could improve it. See Appendix B.10 for details.

Since the coding scheme requires more than commonsense knowledge to understand, the researcher added a short description under the coding examples in Tables D.3 and E.3. For example: an explanation for a "gap identification" process reads “need data for reasons why people chew gum”. Both users rated the case report as describing the experience they went through “very accurately”.

Chapter 4: Processes and Activities

This chapter reports findings about the sensemaking processes and activities that sensemakers engaged in. As this chapter demonstrates, the iterative model proposed in Section 2.5 provides a useful framework for analyzing these activities. In addition to detailed illustrations of how the sensemakers in this study went through each process, this chapter also reports common patterns among different sensemakers as to how they proceed from activity to activity within the iterations.

Section 4.1 illustrates the overall process that sensemakers went through with detailed examples, from gap identification to creating products.

Section 4.2 discusses in detail the paths that sensemakers' processes moved through, including the different patterns of paths, the phases of sensemaking, and the starting and ending points of the iterations.

4.1 Overall Process

4.1.1 Task Planning and Analysis

In several cases, sensemakers started the task session by analyzing the descriptions of their tasks. For example, User MB6 started by reading the case material and figuring what she needed to do for the case analysis:

"OK. So I need to research different demographic groups… So some of the markets that I had researched before will be good demographic groups. Most of them are pretty equal, I think, if I remember correctly... Let's see, yeah …either white or

Asian. Um... but they're all pretty equal, so I guess I'll just research a couple of them... Let me see... Occupations, business..." (User MB9)

Often previous knowledge was brought in when doing task analysis (*Most of them are pretty equal, I think, if I remember correctly...*). Sensemakers sometimes set up goals and divided the task into sub-tasks.

Some users also planed approximately how long they were to spend for sensemaking and writing up the report/story. For example:

"I have 2 hours and 40 minutes, and I don't want to run out of time… I should use half that time, 1 hour maybe, researching, 1 hour and a half, writing." (User MJ9)

Some users started their task session without explicit task analysis and planning. For example, User MJ3 started by saying *"I'm just going to do a Google search."* Some users skipped task analysis because they have previous worked on the task (mentioned in the think-aloud protocols). Some users seemed to skip task analysis when the task is relatively simple and familiar. For example, the post-session interview with User MJ13 suggested that she had written a similar story on energy in the summer before the semester during which the study was conducted, and that was very likely the reason why she started a Google search directly.

4.1.2 Gap Identification: Reaching the Unknown from the Known

Gap identification or gap definition, often considered to initiate other sensemaking activities to "bridge the gaps" (Dervin 1980), is ideally the first step for any sensemaking tasks. However, as observed by other researchers (Klein, Moon et al. 2006), in some sensemaking scenarios (for example, expert decision making), sensemakers may already know what their information need is and skip this step and go directly into search or building structure. The cases within this study exhibited sensemaking both with and without initial gap identification. Moreover, gap identification also happened at other stages of sensemaking besides the initial stage and was achieved in several ways.

4.1.2.1 Different types of gaps

Ideally, at any point of the sensemaking, there is a certain *set of relevant knowledge* needed to proceed or to claim accomplishment of the sensemaking. A sensemaker may know exactly what the set is, or she may need to discover it as she proceeds with the sensemaking task. The set of relevant knowledge may be static for structured or well-defined tasks, or it may be dynamic or user-defined as the sensemaker's understanding of the task changes, if the task is less structured or defined.

Assuming there is such a set (static or dynamic), gap identification is actually a two-step process in which the sensemaker first identifies this "needed" knowledge separated out from the general knowledge mass and then assesses what she knows / does not know about this set.

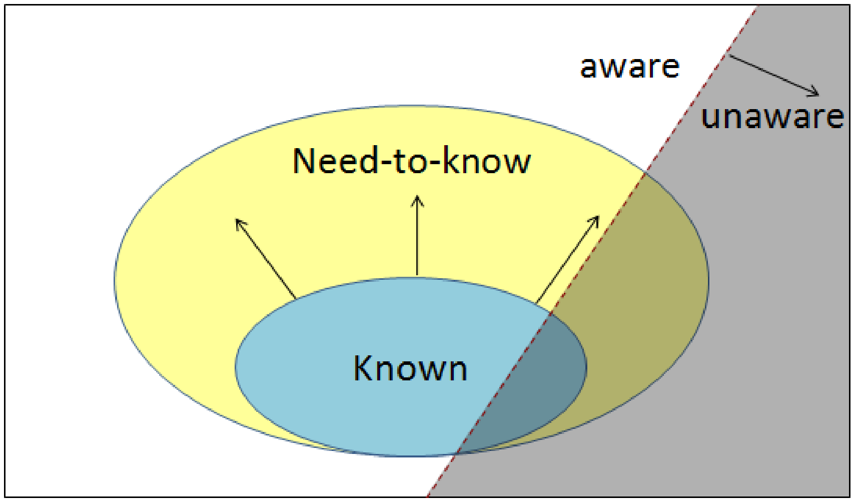


Figure 4-1: Gap Identification

The large oval represents the set of knowledge a sensemaker needs to know for a sensemaking task. The small oval inside it represents what a sensemaker knows about the set. A sensemaker may not be fully aware of what she knows and what needs to be known, and the shaded part represents what the sensemaker is unaware of. Gap identification is the process that identifies the boundary of the known and unknown.

The goal of sensemaking is to move in the directions of the arrows, to enlarge the known part by filling in the need-to-know part as far as possible and to increase awareness of both parts.

4.1.2.2 Data Gap and Structure Gap

Given the assumption that knowledge is stored in the brain as schemata, i.e., structures instantiated with data, gaps in knowledge can be divided into structure gaps and data gaps.

Data gap. In some cases, the sensemaker's knowledge structure was well-formed or at least satisfactory, but she identified a lack of specific data pieces. For example, user MB5 had constructed his “potential markets” concept which she further divided into “youth market” and “adult market”. The user identified a data gap about the “youth market”:

*“OK, I need some **facts** about the youth. So ‘Americans ages eighteen to twenty-four…’ some **facts** about that.”*

The user identified another data gap as she went into “adult market”:

“… I start going into the adults. OK, so I think I just, I mean, I definitely have **something solid**… Um, I need to add **the facts** and stuff.”

Sometimes a data gap can be very specific, for example: *“I would really like to find some **polls** on other energy issues especially **on the price of gasoline** because that has been beat into the ground.”* User MJ13, working on the Energy News task, identified a data gap which is “polls on the price of gasoline”. These specific data gaps tend to be recognized in the later stages of sensemaking where the general understanding (structure) is constructed, and specific facts are needed to fill in the holes. For example, a user working on the new story needed *“a poll that says x amount of people care about energy.”* At the report-writing stage, it can be even more specific, mainly sought to be used as evidence or quotes to support the statements that sensemakers make: *“I need somewhere where it says that, I mean, besides just sales, that Orbit is considerably better.”*

The needed data has become very specific not only in content, but sometimes also in the desired formats.

*“I need just, like, **a number** that says teens are chewing lots and lots of gum.”*

*“I want a **table** that shows me why people are chewing gum.”*

Sometimes there were other restrictions on the desired data. For example, *"Anyway, it's like, two months old, but we'll see if I can find anything newer." (MJ4)* Being current is one of the most common and important quality of the data especially for the news writing task.

Structure gap. Sometimes the identified gap has to do with structure: what the involved concepts are and how they relate to each other. Often structure gaps seemed a lot more "unclear" to sensemakers, compared to data gaps where the sensemakers knew what was missing. In these cases, sensemakers know they do not quite understand something, but they do not know exactly what is there to understand.

"I guess I did not know too much about the equipment rental business so I don't know if my factors are necessarily true." (User MB4)

"I think I do not really understand completely about offshore drilling." (User MJ13)

In the above examples, the sensemakers realized that they lack knowledge about some subjects (equipment rental business or offshore drilling). The lack of understanding seemed to be at both structure and data levels, but the structure gap seemed to dominate. Not until the structures are built can the sensemakers attempt to further identify any data gaps.

Sometimes the structure gaps can be specific about the concepts or relationships. For example, *"Hah, that's good detail that I don't know how to tie in though" (user MJ3)* indicates that she did not quite understand how a specific piece of data relates to other parts of her knowledge structure.

4.1.2.3 Ways of Gap Identification

Gap identification by task analysis. Gaps can be identified by analyzing the task requirements. For example, user MB5, working on developing a marketing proposal for a gum product, identified the set of "needed" information at the beginning of the task session. "*So my main, main task is for me to research Trident gum. Um. Have to look at pretty much **the whole product**, **the advertising that they currently do**…the, the advertising that, **the marketing that they currently do**, um, and then look at some of **the competitors**, um, research **the market itself**, um, kind of see what's out there, what other people are doing, um, different stuff like that.*" By analyzing what is needed for the task the sensemaker identified several issues to research about, including

- "the whole product (gum)",
- "the advertising (marketing) that they (Trident) currently do",
- "the competitors", and
- "the market itself".

This was the set of relevant knowledge needed for the task, which also seemed to be what he did not know much about at this beginning stage.

Gap identification by questioning / Socratic dialogue. Some users conducted Socratic dialogue to examine the completeness, consistence, and anomalies of their knowledge. For example, user MB5 asked a sequence of questions to identify the gap in his knowledge about the problems that Trident has in its marketing:

"But what are they doing exactly? And what are problems specifically with Trident?"

"So the problem is that Trident has… um, competitors… <sp /> so, let's see, Orbit…

What is it? What did they say?" "What other problems did they have? <sp /> I guess the problem is that they're seen as too traditional, um…by kids?"

By questioning, user MB5 was able to identify the gaps that he needed to fill in his knowledge, i.e., the problems that Trident needed to address; the questions provided a useful guide for his search and sensemaking.

These gap-identifying questions sometimes appeared as part of the sensemakers' thinking process (i.e., appeared in the think-aloud protocol). Sometimes they were written down in the workspace as notes to indicate need for further investigation or thinking, as user MB4 mentioned in the post-session interview: *"… I made questions that I am going to answer under that section (recommendation). Like what's the risk? So just things to think about later."* Some other examples of the questions she wrote down include: *"What should the company do to improve its net earnings?" "How will they finance new growth?"*

Gap identification by experience. Expert users may be able to identify gaps based on their previous experience working with similar tasks. For example, when user MJ4 started writing her news story on energy and election, she said *"And then I need the quotes. Just because you can sense when, when you need a quote in a story."* The user had training in writing news stories, and she knew when she would need a quote naturally in the work flow.

Ad-hoc gap identification. Some gaps were identified not through a purposeful act but as a byproduct of other activities. A sensemaker might find articles about one concept, read them, and encounter a new concept that he/she did not understand. He or she would naturally ask the question, "What is it?"

Sometimes the newly discovered concept could be important to understand first so that the other concept could be comprehended; sometimes the newly discovered concept did not have much to do with the central task and the sensemaker was merely asking the question out of curiosity. These curiosity-triggered gaps often did not get filled, although occasionally some sensemakers did go find the information to satisfy their curiosity.

4.1.2.4 Dealing with Gaps

Often the identification of a gap triggers some further action, either searching for information to fill the gap or recording it in some other way so that one can keep track of it and deal with it later.

Search for information immediately. For gaps that were considered critical, sensemakers immediately went searching for information to close the gap. Sometimes a gap stopped the ongoing iteration of search and sensemaking. Sometimes the sensemaker came back and continued with the paused iteration when the gap was filled. Sometimes the gap altered the direction and started new search-sensemaking loops.

For example, when reading about an article about tax credits, user MJ14 encountered a term "Alternative Minimum Tax patch", which initiated a list of questions: *"What's this 'Alternative Minimum Tax patch'? What is that? Let me just look that up really quick."* She selected and copied "Alternative Minimum Tax patch" from the article, and pasted the phrase into the Google search box to start a new search.

Search for information **in the next iteration.** If the identified gap was not perceived as so critical that the user had to stop the current iteration, the user finished the current iteration first and then started a new iteration to search for information to bridge the gap.

Record the gap and move on. Sensemakers simply acknowledged that there was a lack of understanding. This can be done by writing questions on a separate note page, or it can be done by inserting a "place holder" in the writing. For example, while working on an advertising case for Trident gum in the first session, a user created a list of "Continue to Look Into" in his case write-up in which he included items that needed further research.

Ignore the gap. There are also cases where gaps may be recognized but never get dealt with. For example, when user MJ3 noticed that she did not know about a bill that she saw, she asked, *"What is this? 'The Warner-Lieberman Bill.' "* She recognized a gap in her knowledge about this particular bill. However, this was as far as she got in regards to the gap. She did not make any effort to "bridge" the gap, simply because it was not important enough to the task.

4.1.3 Search

Search is often the next step following gap identification (although some gaps never get dealt with as discussed in the previous section). Depending on whether a search has preset goals and strategies, a search can be classified as exploratory or focused search. The cases in this study showed both exploratory and focused search. Another facet of search is whether the sensemaker searches for data or structure. In many cases, focused search followed exploratory search; sometimes focused searches were embedded in exploratory search. The distinction of whether a search was for data or structure is not always clear. Sometimes a sensemaker was undetermined, especially with exploratory search, as to whether what he looked for was structure or data.

4.1.3.1 Exploratory Search

In many cases, a sensemaker started with little knowledge about the subject area. The goal of the first few rounds of sensemaking is often to learn rather than to find specific information (Marchionini 2006). For example, a sensemaker working on an advertisement proposal learned about the current status of a product's marketing effort before he went searching for any particular ad or campaign. Then the search became exploratory again when he started *investigating* the options that the product had in terms of improving its market share.

Exploratory Search for Structure

Structural gap seemed more likely to be followed by exploratory search, since the information need for structure is difficult to express using traditional keyword queries (Qu and Furnas 2007). Since a structural gap often co-occurred with a data gap resulting in a fuzzy unclearness about a concept or category, the sensemakers did not necessarily have in mind a decision to explore a concept for structure rather than data, or vice versa.

Browsing a familiar source was an effective approach to get some structures. With the *Energy News* task, sensemakers often used news websites where they checked for news regularly, such as washingtonpost.com, nytimes.com. For example, a sensemaker (MJ15) looking for *"information about what are the main aspects of* *the energy in the US"* went to a candidate's Web site directly to browse the website. She put the mouse over tabs "PEOPLE", "ACTION", "MEDIA", and "ISSUES", clicked on "ISSUES → Energy & Environment", then browsed the webpage, clicking on the link "Read the full version of The Obama-Biden New Energy for America plan", and then browsing the pdf file in Internet Explorer.

Often these websites' organization of various issues provided explicit structural representation. For example, it might list several issues under *Energy* in the *2008 Election* section. Such websites usually provided an overview of various issues which often gave an implicit structure that was much easier to extract than other relevant documents. With a marketing case, the sensemaker started with a business database (Mintel), and browsed the relevant sectors in a similar way.

Another effective way to acquire structure was to browse the headings of a good relevant document (this was particularly the case when, as with some business cases, the printed case materials were given to the sensemakers and they did not need to do any searching). The headings provided a good structure of the problem to be analyzed. Some users actually adopted the structure of the case descriptions as part of their knowledge structure for the sensemaking task.

Exploratory search for structure did not always succeed because of the uncontrolled nature of the "structureness" of resources. For example, a user (MJ4) browsed both candidates' Web sites, trying to find some structure on the energy and environment related issues. However, neither candidate imposed any structure on the issue. *"... Obama doesn't have- this is listed in alphabetical order, which I think is interesting... whereas McCain's 'Issues' are kind of random. Strange."*

Exploratory Search for Data

Although it was probably not effective to look for a particular fact through browsing, facts or data can also be acquired in exploratory search. A few patterns as to how sensemakers acquired data through exploratory search emerged from the analysis of screen activities along with the think-aloud protocol.

Many sensemakers first used a very broad search term to get whatever they could get, and immediately started browsing the results, not looking for anything in particular, but to get a general idea of what information was available. They collected whatever facts seemed relevant to the overall task, parked them as notes in a general section (sometimes titled *problem*, *overview*, or *background*). Some of the facts that gave a general understanding of the problem or situation remained there; others were distributed to proper sections of their notes later on.

Some sensemakers started with a familiar Web site or database, such as cnn.com, gallup.com, washingtonpost.com, LexisNexis, Mintel (a business database), browsed the Web site, and collected facts in the same fashion. These facts were often collected as a byproduct of trying to get the "big picture" where structure was the primary goal of the exploration.

4.1.3.2 Focused Search

Focused search often followed a data gap or a structure gap that was well defined. In these circumstances, the search actions were very much directed toward answering the questions that were raised in the gap identification step.

Focused Search for Data

Keyword search using a Web search engine or a database is the most common means to find specific data. In focused search, the keywords become much more specific. For example, a sensemaker (MB5) used trident gum ingredient to search for a special whitening ingredient that trident used in their products. He found the name of the ingredient with this key word search, and searched with keywords Xylitol trident to get further facts, such as "*Xylitol helps fight cavities.*"

Known item search is another example of focused search for data. For example, a sensemaker remembered reading a commentary about the two candidates' positions on different issues. First he used words and phrases he remembered from the article for a Google search, but failed to locate the article. He then revisited several places where he might have seen it, and ended up finding the article in his email.

Although browsing is usually used in exploratory search, sometimes sensemakers browsed an article known to be relevant (either assigned to the sensemakers or found and judged relevant by them earlier) with a particular information need kept in mind. Compared to exploratory browsing, focused browsing was more like scanning or filtering, and the sensemakers picked only information that satisfied the focused information need.

For another example, consider User MB6, she looked for demographics of several ethnic groups to develop marketing plans for a mobile phone company; she went to a known database and clicked on the menu sidebars and links *"Basic Counts/Population"* → *"Total Population"* → *"Race and Ethnicity"* on the census bureau Web site. This way she was able to accurately locate the sizes of each demographic group of her interest, which might not have been easily found through key word search.

Focused Search for Structure

As mentioned above, search for structure was usually conducted in an exploratory fashion. Focused search for structure was difficult to conduct using conventional keyword search. For example, a sensemaker looking for "major aspects of the energy issue" used key words "energy issues in the us" to search the EBSCO database at a University database portal, which did not find any results. She changed the key words to "energy crisis" and found "an overview of the different energy problems in the US".

Most structures were acquired using browsing, either focused or exploratory. Some structures were also acquired by serendipity or as a by-product when looking for other information.

4.1.4 Building Structure

Building structure is an essential process in sensemaking. Structures may be built by adapting and modifying others' structures, or they may be built inductively from new information. The built structure provides the framework of the understanding created for sensemaking

Adapting Others' Structure

Sometimes sensemakers were able to find or extract structures established by others. For example, when user MJ3 examined the NY Times election 2008 site on the energy issue, she found good structure segments. The Web site had listed several sub-issues of the overall energy issue, including:

- Federal gas tax holiday
- Taxing oil company windfall profits
- Domestic drilling
- Expanding nuclear power
- Coal plants and coal-to-liquid fuel
- Ethanol subsidies

The user adapted the structure, copied these sub-issues into her notes, and later put them into her concept map. This structure became a major part of her knowledge structure.

Adopting others' structure seemed to be one of the most efficient ways to build structures. Such cases happen when the sensemaker was at an open-minded stage. Some structures were found through purposeful search, and others were acquired by serendipity.

Building Structure by Task Analysis

Some structure elements are built from analyzing the task requirements and what the sensemaker needs to know to complete the task. Task analysis often helped the sensemaker to elicit an initial structure as the starting point for sensemaking. For example, by thinking about what needed to be known in order to formulate an advertisement proposal for Trident gum, a sensemaker (MB5) elicited three aspects of the marketing research:

- The advertising that Trident currently does
- Competitors
- The market itself

These three aspects became the three initial structural element of the sensemaker's knowledge frame (represented as OneNote pages): Trident, Gum Sector, and Competitors. The initial structure was expanded and modified later on when the sensemaker did more research.

Building Structure from Prior Knowledge

Prior knowledge also contributes to structure building. It is rare in real life sensemaking tasks that the sensemaker does not know anything about the domain or the task. In the cases of the study, the sensemaker always had at least some prior knowledge. The sensemaker might not have known some concepts and relationships in the task domain, or she may have had some general knowledge about how to approach a task. Even if the initial structure was far from satisfactory, it was sometimes more efficient to build on the concepts and relationships in the initial structure than to build everything from scratch.

It seemed like some concepts, at least the key concepts, were often known to the sensemakers. For example, the broad categories representing major aspects of the task domain such as candidates of the presidential election or major issues of a business plan were often created right at the beginning before any search activities. But the concepts may not have been connected in a way that helped accomplish the task at hand, so relationships were constructed on the fly to address the need of the task. For example, a sensemaker may have known McCain and Obama were two candidates of the 2008 presidential election, but not necessarily how the candidates compared to each other in regards to their energy policies.

General knowledge about the task or topic also contributed to structure building. Most users working on the energy news assignment created two nodes for the two candidates in the election, not based on what they found, but rather on their general knowledge of the presidential election.

Building Structure from New Information

Sometimes no existing structure is available, ready to be adopted. When this happened, the sensemakers needed to construct structures from scratch. They consumed the new information found, generalized it to a conceptual level, and fit the newly built structure element into their prior knowledge.

For example, by reading an article from *Time* about the economic downturn at the time of the study, titled *"Will the Environment Lose Out to the Economy?"* , the sensemaker came up with "the economy defeating the environment" concept, which she included in her concept map as an important factor for the news story. This concept was not mentioned in the task description, nor did the sensemaker come in knowing this was going to be what her story would be about; from reading the article, the sensemaker decided it was important for her to *"address this… economy… issue of the economy defeating the environment…."* This concept came from a section in the article which provided an idea she was to use in her own story: *"with the tanking economy dominating the news, and the government willing to virtually bankrupt itself to bail out the financial sector, it could be hard to push the climate change agenda – and possibly hard to find any money left to support it."*

New information also helps the consolidation of prior knowledge. In many cases, sensemakers may have known about certain concepts, but they waited until they found information about these concepts to create note pages or nodes for these concepts. For example, a sensemaker knew that Obama and McCain were the two candidates to be discussed for their energy policies, but she did not create note pages until she located and browsed through their campaign Web sites on energy and environment issues. (*Example MJ4*). In these cases, new information helped the externalization of prior knowledge and sometimes helped the construction of previously known concepts into a structure that was relevant to the sensemaking task.

Many cognitive mechanisms were used by sensemakers while they were trying to build structures, to discover and eliminate concepts and connect them in a meaningful way that was compatible with their existing knowledge. See Chapter 6 for findings on cognitive mechanisms.

4.1.5 Instantiating Structure

Note taking was the most common way of instantiating structures in this study. Facts were added to different note pages as instantiation of the structure. Notes were taken in forms of copied notes, restated notes, and summary notes. Instantiation of structure always led to accretion, the factual addition of knowledge. See more discussion about accretion in Section 6.1.1.

Another way to instantiate a structure is to link notes in OneNote to a concept node in CMap. This can be done by adding a hyperlink to the CMap concept node. Clicking on the hyperlink will open and highlight the linked notes in OneNote. For example, User MB6 created nodes labeled with populations of ethnic groups of interest in CMap, and then linked to the numbers in her notes copied from the census bureau with detailed information on that. This approach allows different levels of abstraction to be seen through multiple presentations.

Sometimes sensemakers discovered facts that were potentially useful to the task, but had difficulty fitting them into the current knowledge structure. Sometimes this difficulty caused residue which led to representation shift (Russell 1993). Sometimes the unfitted facts did not necessarily result in restructuring. For example, a sensemaker may not have necessarily been ready to create new structures to accommodate the residues, or she may have had more important mini-tasks at hand and needed to come back to these facts later. In these cases, the unsorted notes were put in a separate note page--sometimes untitled, sometimes given very general labels such as “general info”, “problem”, or “background”. This created a temporary storage space to “park” notes that did not quite fit into the existing categories. These semi-instantiated notes sometimes turned out to be not useful after all; sometimes they turned out to be crucial to the restructuring of knowledge structure.

Sometimes sensemakers stored facts that were potentially useful. For example, a user discovered a detailed energy plan from Obama’s Web site but was not sure if she would use it at all. She still copied it into her notes. *“I doubt that I’ll use it, but it couldn’t hurt. This could give me more, more details, probably; if I find a bullet point I don’t understand or want more info on. But other than that, I don’t think I would be using this.”*

Most of the time, instantiating structure happened once a structure element was built; sometimes it happened synchronously with building structure. See section 5.2 for more details on the relationship between the two processes.

4.1.6 Consuming Knowledge Structure and Creating Work Products

Knowledge was updated throughout the sensemaking process. And the final work product, in the format of a report, news story, presentation outline and notes, or answers to case questions, was the final reflection of these continuous updates. The creation of the work product often relied on the knowledge structures created and instantiated. Usually all instantiated structures had to be consumed by the sensemaker to create the work product. There were also cases where new structures and data were sought during the process of writing the story or report.

Most of the time the process of consumption and knowledge updates was highly internal and was not always reflected by traceable activities of the sensemaker. However, in some cases, there were concluding periods in different phases of a sensemaking task reflected by the think-aloud protocols. For example, when a sensemaker finished researching the polling results on energy, she stopped other activities and used a minute or two to go over the major findings again in her mind to update her understanding of public opinions as reflected by the polling results. In another example, a sensemaker went through reflective thought after each section she read in an article, recapitulating what had been read, and connecting this to prior knowledge. Through periodically stopping in order to think, sensemakers were able to consume the newly acquired knowledge (structure and/or data) and update their existing knowledge accordingly.

The final product of sensemaking was not merely a collection of accumulated knowledge updates. In some tasks, it was marked by the creation of new knowledge, such as a marketing plan that included not only the structure and data that were acquired or constructed during the search and sensemaking process, but also creative advertisement ideas that were crucial to the success of the sensemaking. Such ideas may be inspired when the sensemaker went through constructing his conceptual space, but they were more carefully crafted and elaborated on how they were to be put to action. In the energy news examples, sometimes insightful opinions were formed at the final stage as a sensemaker wrote his story. He may have used data found and quotes from other people, but the opinions he infused in his story were original and not found somewhere else. In both examples, the knowledge updates in the sensemakers conceptual space did not happen by just accumulating knowledge and connecting it to prior knowledge; it involved some creativity (Klein, Moon et al. 2006) which is one of the most important aspects of sensemaking.

4.2 Paths

Examining the paths that sensemakers went through showed interesting patterns of paths that are useful for designing system to assist users with their sensemaking tasks. This user study confirms that the simplified waterfall model of how data lead to knowledge and understanding runs counter to empirical evidence, as observed in (Stefik, Baldonado et al. 1999; Klein, Moon et al. 2006). In fact, the paths that sensemakers went through in this study were rather untidy and heterogeneous. However, as idiosyncratic as the paths taken by individual sensemakers were, there were some patterns that appeared across many cases. The model proposed in chapter 2 provides a useful analytical framework for looking at these paths.

Section 4.2.1 reports findings about patterns of the paths.

Section 4.2.2 describes the phases of sensemaking that constituted several iterations.

Section 4.2.3 discusses the starting and ending points of sensemaking iterations.

4.2.1 Patterns of Paths

As the model in Chapter 2 suggested, the whole sensemaking process may be divided into several “search—sensemaking” iterations. The analysis revealed a large spectrum from planned, systematic to rather random, ad hoc patterns of the paths of the search-sensemaking iterations. This section reports these paths with examples. There were slight variations of paths shown in each of the reported categories, and the examples here are not intended to be inclusive of all but rather to give concrete illustrations of what a typical path in each category would look like.

4.2.1.1 A Systematic Path

In some iterations, a sensemaker went through a full-fledged path that went from task analysis, gap identification, search, to building structure and instantiating structure, especially at the beginning of a sensemaking task or after a failed iteration a sensemaker decided to “start over”. There were slight variations in terms of the exact order of these processes. Figure 4-2 shows the path taken by user MB5 in his first iteration.

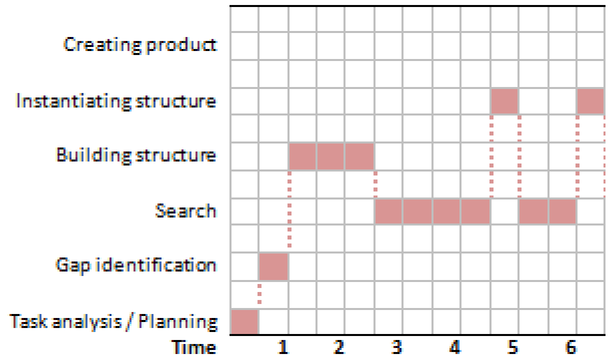


Figure 4-2: A Systematic Path by User MB5

In this iteration, user MB5 started analyzing his task, identified gaps in his knowledge which became his initial structure, and searched for information to instantiate the initial structure.

Figure 4-3 shows the path user MJ13 took during her first iteration.

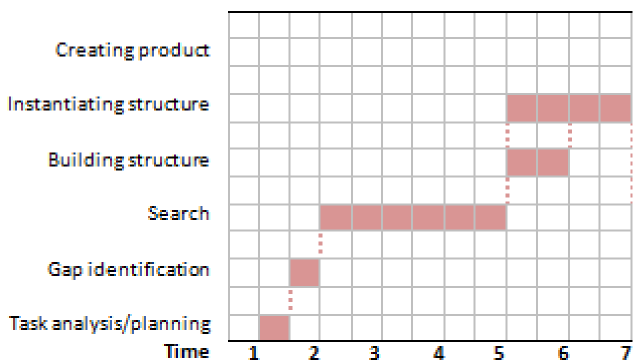


Figure 4-3: A Systematic Path by User MJ13

In this example, user MJ13 started with task analysis, identified gaps in her knowledge, and started to search for the missing knowledge immediately; she found a useful article, built a structure from it, and instantiated the structure simultaneously.

The first example illustrated a systematic sensemaking approach that is top-down, or structure driven. The second example illustrated a systematic sensemaking approach that is bottom-up, or data-driven. See Chapter 7 for more details on top-down and bottom-up sensemaking approaches.

4.2.1.2 Gap Identification – Search – Instantiating Structure

In many cases, sensemakers skipped task analysis since it was not necessary for all iterations. Most users analyzed their tasks at the very beginning, possibly also reanalyzing it somewhere midway through. Several iterations started with gap identification, which led to search and building or instantiating structure. The following figure shows an example of a "gap identification – search – instantiating structure" path:

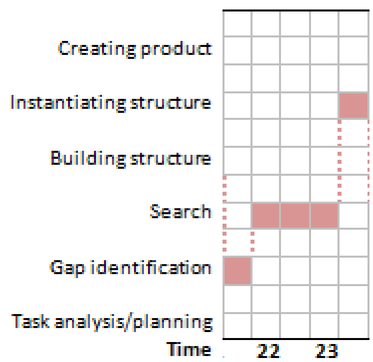


Figure 4-4: "Gap Identification – Search – Instantiating Structure" by User MB5

In this example, user MB5 identified a gap in his knowledge; he needed *"some Trident-specific marketing articles"*, so he did a Google Web keyword search with "trident gum marketing", found an article titled "Gum Wars", read the article, and copied and pasted paragraphs about Wrigley's new marketing efforts into OneNote in the "Trident" page.

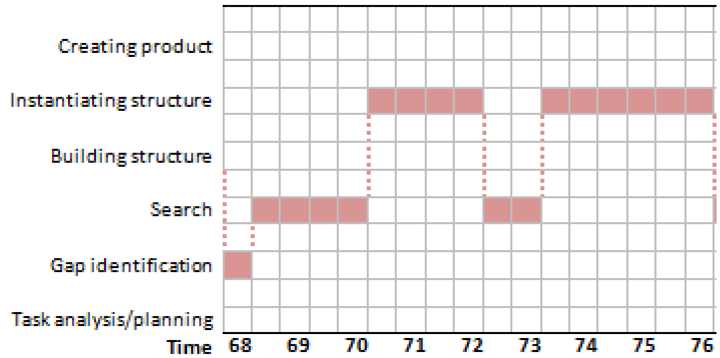


Figure 4-5: "Gap Identification – Search – Instantiating Structure" by User MJ15

In this example, User MJ15 decided that she needed some information about "oil prices" which was one of the issues she decided to talk about in her article. She did a Google Web search, found a relevant article, and put it in her notes. Then she continued to examine the results list and found another useful article and put it in her notes as well. Note that the sensemaker went through two consecutive "search – instantiating structure" sequences which had similar goals and outcomes.

As shown in the above examples, the "gap identification – search – instantiating structure" path was most common when a satisfactory structure was present. They tended to occur in consecutive sequence (shown in Figure 4-5) until the sensemaker felt that enough facts were accumulated to move on.

4.2.1.3 Search – Building – Instantiating

In some cases, gap identification was not explicitly present. Sometimes gaps were identified in earlier iterations; sometimes the sensemaker did not take the effort to recognize what was missing in her knowledge and moved directly into searching.

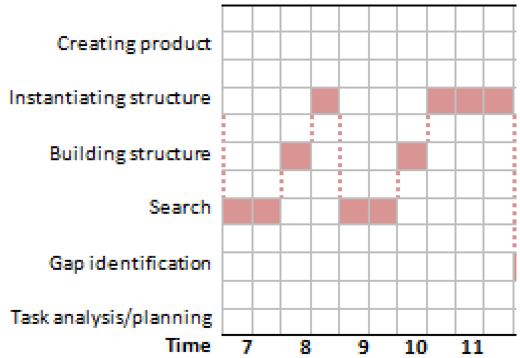


Figure 4-6: "Search – Building – Instantiating" by User MB5

In this example, user MB5 was scanning a report that he found relevant. After scanning through the sub-section Advertising and Promotion, he created a note page titled "Advertisements", and copied and pasted the summary paragraph in that section to his notes. Then he continued to scan the "Customers" section of the report, created a note page titled "Problem/Opportunity", and copied and pasted several facts on youths' and adults' use of gum.

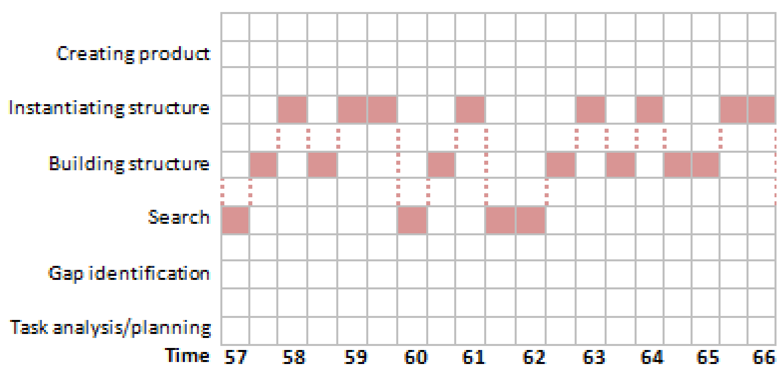


Figure 4-7: "Search – Building – Instantiating" by User MJ9

In this example, the sensemaker was visualizing public opinion in CMap. She looked through her notes on polling results in OneNote, created a node in CMap ("Nuclear Energy"), and typed in "Mixed Reaction 47 for 41 against" before creating another node, "Offshore drilling", and typing in "Repubs and Independents". She continued to do this for other issues including "Conservation", "Windfall Tax", and so on.

This type of path is a typical abridged iteration. Successful search found useful materials which lead to building a structure piece and instantiating it immediately after it was built.

4.2.1.4 Search – Instantiating – Building

Sometimes, usually when some initial or intermediate structure was built, search followed by instantiating the existing structure resulted in building of new structure. Sometimes sub-concepts emerged within the categories instantiated, and sometimes new concepts and relationships were discovered that were at the same level as the concepts instantiated.

In the following example, the sensemaker (MB5) found some information about different uses of gum by youth and adult audiences. He put them under "Problem/Opportunity", showing that he recognized some common themes of the notes under this page, including bold flavors and caffeinated gums for teens and whitening for older adults. He then created separate note boxes for these themes (building structure), and then copied more information about each (instantiating more). Then he created a few nodes in the Concept map to represent these concepts (building structure visually).

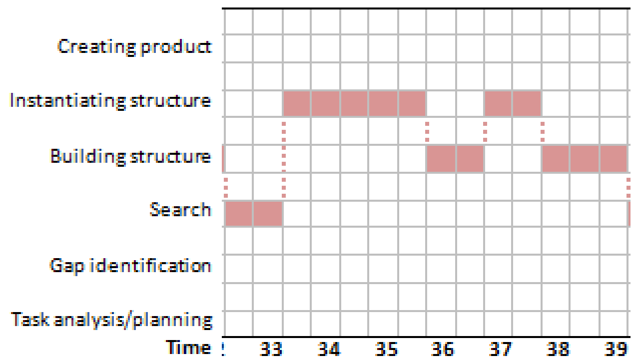


Figure 4-8: "Search – Instantiating – Building" by User MB5

This seemed to be an effective way for the sensemaker to recognize patterns in the new information (different patterns in gum use by younger and older audiences), see the connection of these use patterns as opportunities for marketing, and then transform them into a structure (youth and adult market segments) that was intended for the task output.

In the second example (shown in Figure 4-9), user MJ15 looked for global warming and sustainable energy, found a poll (search), and put it in her notes (instantiating structure). She continued to search and found an article talking about sustainable energy and put in notes (instantiating structure). At this point she decided that she was not going to talk about global warming (building structure): *"Basically I am not going to talk about global warming. But I am going to talk about sustainable energy and the benefit of it."*

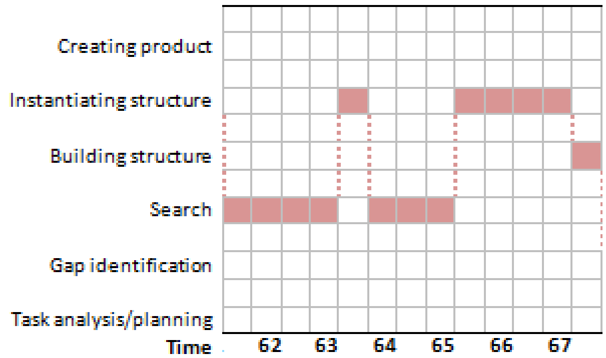


Figure 4-9 "Search – Instantiating – Building" by User MJ15

The modification of structure (talking about sustainable energy and the benefit of it instead of global warming which was originally planned) was based on the new information found. Instantiating structures with the new information made the sensemaker aware that there was not much useful information about global warming she could possibly use in her story, and that led to an action of building structure.

This pattern often suggested a combined top-down and bottom-up approach, where the sensemaker started top down to search for information to fit an existing structure, and went on to discover any theme or structure that appeared in the new material (bottom-up), and then tried to connect the newly discovered structure which may not have entirely met the requirements of the task.

4.2.1.5 Search – Building Structure

In the following example (Figure 4-10), the sensemaker searched his notes, which were somewhat structured with several note pages and implicit structure embedded in each page. He searched for structure in the semi-structured notes and used concept mapping tools to build structures or make implicit structures explicit.

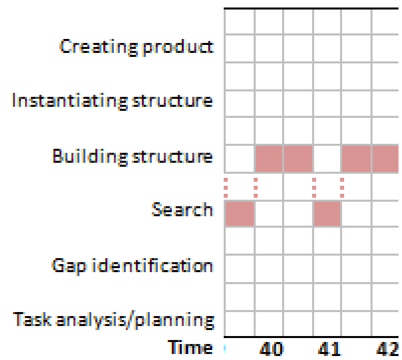


Figure 4-10: "Search – Building Structure" by User MB5

Although sometimes requiring extra time and consideration, the effort to make implicit structure explicit or even to represent and view the already existing structure in a different perspective (for example, creating a concept map based on notes) seemed to help the sensemaker to "sort out the ideas a little better". For example, with this task, it was during this iteration that the idea of having two main campaigns--one focusing on the younger audience, including bold flavor and selling the bold flavors, and one for the older audiences, selling the health aspect and the whitening. *"Thinking, though, that if they could combine it, it could be really, really powerful. Um, you know, sell the, the health to the adults, and the, the coolness factor to the kids."*

This pattern suggested a top-down approach where the sensemaker was structure-driven. Searching for structure followed by building structure including making implicit structure explicit often constituted critical phases where the "meaning" was constructed. Scattered pieces of information were put together in a connected way which also connected to existing knowledge of the sensemaker.

4.2.1.6 Search – Instantiating Structure

"Search - Instantiating Structure" is one of the most common and frequent patterns. It happened when a relatively stable structure was built, and the sensemaker simply searched for data to instantiate the structure. This pattern often occurred multiple times until the sensemaker found enough data about to move on to the next iteration.

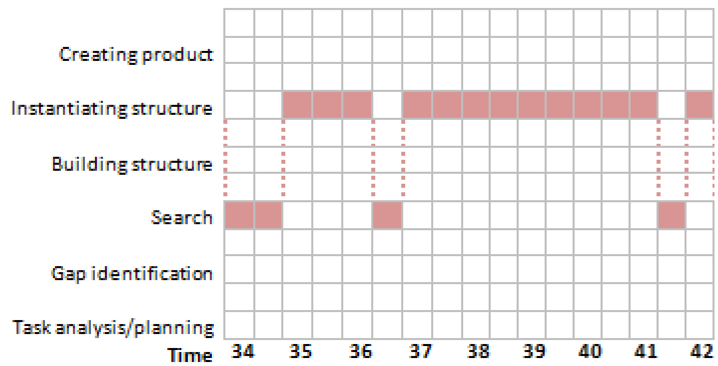


Figure 4-11: "Search – Instantiating Structure" by User MJ9

In this example, user MJ9 had established a stable structure including a note page for "polling results". She searched the Gallup Web site to locate polling information on energy issue. She found several polls and put them in her notes one at a time.

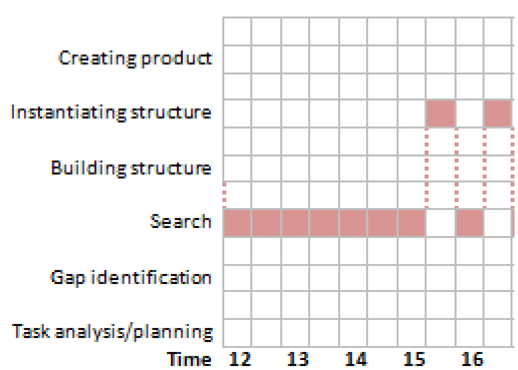


Figure 4-12: "Search – Instantiating Structure" by User MB1

User MB1 had case materials assigned to him; he browsed the case materials (search) to select useful pieces of information, added a few bullets into the his notes, and then continued to browse and put in more notes.

4.2.1.7 Building and Instantiating Structure in Sequence

In some cases, the sensemakers did not need to search for information because what they needed to make sense was given to them. In the following example, user MB4, working on a business analysis case, was assigned an article by the instructor of her course. Everything she needed to answer the case questions was included, so she did not do any search to find additional information. She built structure by creating note pages from the sections of the article, and instantiated them with restated notes of the content.

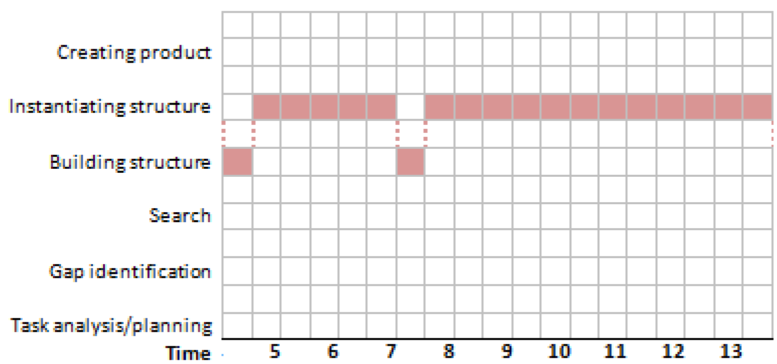


Figure 4-13: "Building – Instantiating Structure" by User MB4

With this pattern, reading and comprehending a given document was the focus. The instantiated structure represented as note pages and/or concept maps were often visual representations of the document, whereas sensemaking tasks that required finding multiple pieces of information and putting them together in a meaningful way often resulted in a representation that reflected the sensemaker's own perspective.

4.2.1.8 Building and Instantiating Structure at the Same Time

In some cases, after finding a relevant document, the sensemaker built and instantiated structure simultaneously. This often happened when the document not only provided good details but also structure elements.

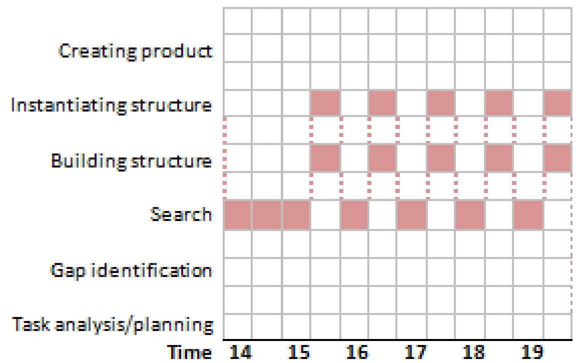


Figure 4-14: "Building and Instantiating Structure Simultaneously" by User MJ5

In this example, User MJ5 searched for Obama's energy policy, and found a useful PDF document from the Website. While browsing that document, which had provided him with both structure and details, the user added notes (instantiating structure) with sub-headings (building structure).

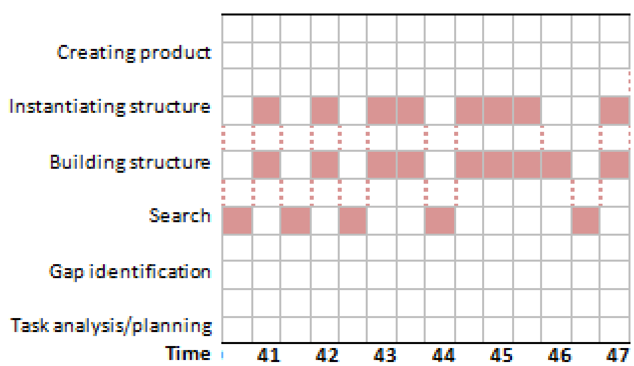


Figure 4-15: "Building and Instantiating Structure Simultaneously" by User MJ4

In the second example, User MJ4 moved to CMap to create a concept map for her story. She searched for ideas to inspire a "lede"[1] from her notes. Then she created a node "lede" and instantiated it with a paragraph. She continued to search through notes, and created and instantiated a node "clear car challenge". She continued to search her notes and created other parts of the map.

In both examples, the information found (either an article that is well structured, or a sensemaker's notes that was organized into pages and sub-headings) presented both structures and data in a meaningful way to the user. When users recognized the instantiated structures, they put both structure and data in their conceptual space simultaneously.

4.2.1.9 "Gap-identification – Search" embedded in Creating Product

While creating the product of sensemaking in the form of a report, news story, or presentation, sensemakers often found new gaps when they needed to present and convey their ideas to others. The gaps were mostly data gaps: sometimes the sensemakers discovered that they needed previously overlooked facts about some concepts, and sometimes the sensemakers needed evidence to support whatever claim they were making. They then went to search for specific facts, and used those facts in their sensemaking product. The facts that were sought at this stage often did not change the sensemakers' established "sense" or understanding; they were rather purposefully chosen to fit or strengthen the understanding.

[1] **lede** (lēd)
n. The introductory portion of a news story, especially the first sentence.
[Obsolete spelling of **lead**[1], revived in modern journalism to distinguish the word from **lead**[2], *strip of metal separating lines of type*.]

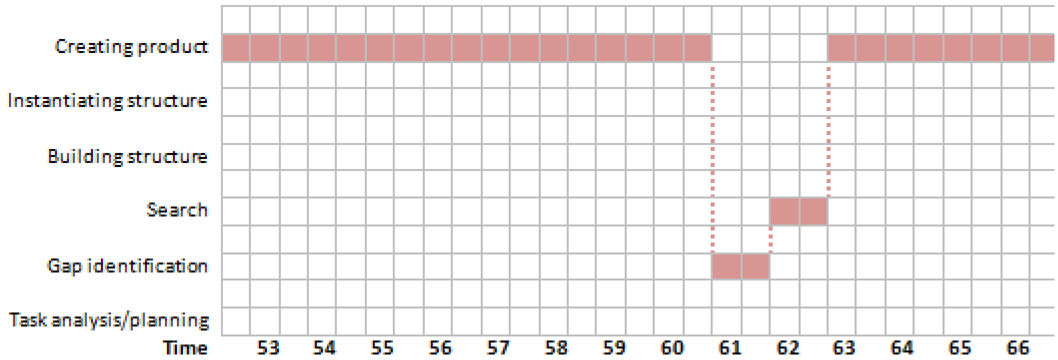


Figure 4-16: "Gap-identification – Search" embedded in Creating Product (MB1)

In this example, while answering a case question about "stock grants" and "stock options" (creating product), User MB1 realized that he needed to learn more about "stock grant"; he looked for specific facts about stock grant and used them in his analysis.

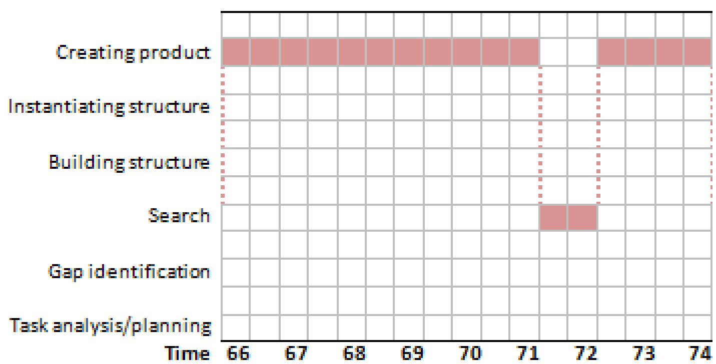


Figure 4-17: "Search" embedded in "Creating Product" (MJ4)

In most cases (such as the second example), "gap identification" was not explicit during the product creation phase. Users just naturally referred to their notes and maps for information to put into the writing without explicitly identifying what was missing. This is partially due to the fact that by the time that sensemakers were ready to write down a report or story, they had already developed a comprehensive picture of the structure they had created and what data was available in their notes.

The purpose of search during this phrase was to "locate" a specific known-item such as a quote they remembered, rather than finding an item that meets a given specification such as a quote by McCain on fuel efficiency. For the same reason, search during product creating often did not result in any instantiating or building of structure since the structure and data was already in the conceptual space.

4.2.2 Phases

4.2.2.1 Exploratory vs. Focused Stages

Search and sensemaking is often exploratory at the beginning, with exceptions when the sensemakers have sufficient knowledge about the topic. For example, a journalism and government major who described himself as "politically intensive" was able to identify what his story was going to focus on and went directly into the focused stage.

In the first few iterations, users explicitly sought "general knowledge" about the topic or issue at hand, looking for "summary" and "overview" to get a "good basis". Users identified structure gaps and sought to bridge them. Users often ignored details and specifics or saved them for later reference. In the focused stage, sensemaking became highly directed by gap identification and bridging. Users extended higher-level structures with more specific concepts. Data gaps were mostly present at this stage, which led to focused search for data, and instantiation of structures with data. Some structure elements may have been abandoned because they did not fit with other parts of the structure.

4.2.2.2 The Crucial Stage of Sensemaking

In several cases, a key point of sensemaking happened about half-way through the task: the dots were connected, a perspective was found, or a solution direction was identified; everything started to “make sense”.

In the energy news example, the key moment happened when the user decided to do an overview story on energy based on what he had learned, instead of focusing on a particular aspect. After that, the search and sensemaking was more directed and the structures became clearer.

In the Trident example, after some research the user came up with a general principle for her advertisement proposal, namely to “sell the health to the adults and the coolness to the kids”. After that, sensemaking became more and more concrete as to the specifics to combine the two factors and what exactly the advertisement should look like.

4.2.3 Starting and Ending Points of Sensemaking

The sensemaking process took place in iterations. Sensemaking can be viewed as a process by which the sensemaker moves through a series of knowledge states, each of these states may be a starting point or ending point of a sensemaking iteration. Several factors may influence how the iterations proceed in terms of how it starts and ends. This section discusses the different situations in which a sensemaker started and exited a sensemaking iteration.

4.2.3.1 Starting a new iteration

There were several reasons for starting new sensemaking iterations. The iterations started either as planned or necessitated by what was found. Planned reasons included identified gaps (explicitly or implicitly), moving on to a different sub-topic, and the temporary completion of a sub-task; perceived anomaly or uncertainty in existing knowledge schema, or the incongruity of the existing knowledge schema in interpreting new information for a given task, are also reasons to initiate a new round of sensemaking.

New gaps identified often triggered a new iteration (See section 4.1.1.3 and 4.1.1.4 for details on different ways of gap identification and how users deal with gaps). Iterations that were triggered by gap identification may have started after a successful sensemaking iteration, in the middle of another sensemaking iteration, or after a failed search.

Different sub-topic. Some sensemakers divided the overall sensemaking topic into several sub-topics. For example, several users tasked with the election news story divided the overall topic "role of energy in the presidential election" into sub-topics such as "Obama's energy policy", "McCain's energy policy", and "Energy polls". When moving to a different sub-topic, sensemakers naturally started a new iteration, which may have been followed by several iterations. New sub-topics may have been all laid out at the beginning or they may have been decided rather randomly along the way. For example, a user (MJ15) decided to research the Democratic Party's positions on energy *"off the top of my head"*.

Different sub-task. Some tasks were divided into sub-tasks. For example, the task of proposing a marketing plan for Trident gum was divided into:

1. Research the gum sector
2. Compare Trident with its competitors

3. Examine existing advertisement effort of Trident

4. Write report

When moving to a different sub-task, regardless of the success or failure of the previous sub-task, the sensemaker naturally started a new iteration, which may have fallen into any of the patterns of paths described in Section 4.2.1.

Uncertainty was often the greatest at the beginning of the sensemaking tasks in the cases under study. Uncertainty may have been associated with the task itself; for example, the question of how to proceed and what was expected at the end may have caused uncertainty, because it may be associated with the knowledge domain, such as the opinions of the presidential candidates. Sensemakers used different strategies to reduce uncertainty. Some analyzed the task to identify gaps in knowledge, or to divide it into several sub-tasks. Some started by exploratory search (including browsing) sources to get general ideas about the domain. The necessity to reduce uncertainty triggered sensemakers to start an iteration from either gap identification, search, or building structure.

Anomaly or incongruity in knowledge was another common trigger for starting a new iteration. The incongruity may exist between what users found in the new material and what was formerly believed by the user, or it may exist between information found from different sources. The conflicts in knowledge schema that sensemakers experienced often required some actions to be completed: either search for more information or find a way to resolve (or not) the conflicts. Anomaly has to be dealt with to get to the goal of sensemaking: a coherent understanding. See Section 5.3.3 for more details on how users deal with conflicts.

4.2.3.2 Ending an iteration

Iterations ended for different reasons, which can be divided into sensemaking success and sensemaking failure. Sensemaking success may be assessed by several measurements, conceptual changes that users achieved being one of them. Moreover, the cases (journalist writing and business analysis) under this study suggested that sensemaking success is also a subjective matter, which may not be the case in other tasks. There might also be different degrees of success. When talking about an iteration ended with success, I did not intend to claim with any numeric measurement that a user had a successful sensemaking in such aspects. When talking about successful sensemaking, I refer to it more generally as the situation in which a sensemaker is able to get something out of the materials found (or drawing conclusions from the absence of certain materials), regardless of the absolute correctness of their claims.

Several factors contributed to the decision of ending an iteration, including:

- Perceived sufficiency in knowledge: the sensemaker felt that enough knowledge has been accumulated to achieve the goals of a task or sub-task
- Reduced uncertainty: although knowledge might still be insufficient, uncertainty felt at the beginning of a task or sub-task was reduced to a level where the sensemaker was comfortable to move on
- Cost-benefit: the cost of continuing to search for more information or trying to figure out some sense with the found material on certain topic

or sub-topic may be too high for the user and outweigh the benefit of getting it done

- Deadline: sensemakers may have to stop a iteration when it is taking too long especially when they work were working under a deadline

The following table summarizes the relationship of the influencing factors and their applicable situation:

Table 4-1: Ending Points of Iterations

Sensemaking / **Search**	**a. Success**	**b. Failure**
1. Useful materials found	1a. Success in analyzing materials resulting in reduced uncertainty or sufficient knowledge	1b. failure to analyze materials due to cost or deadline constraints, or due to lack of background knowledge or thinking ability
2. No useful materials found	2a. Absence of material resulting in reduced uncertainty	2b. Absence of material does not allow conclusions, further search and sensemaking subject to the same constraints as 1b.

1a. Successful search and successful sensemaking. In this case, sensemakers collected useful information from various sources, built a structure or several pieces of structure to hold such information, or successfully put it into existing structure. In many cases, the sensemakers successfully updated their knowledge about the concepts and/or relationships and moved on to the next concepts or relationships. Successful search and sensemaking were almost always marked by conceptual changes, whether the conceptual change was factual addition to the conceptual spaces or changes to the structures.

2a. Empty search, sensemaking success. In some infrequent cases, failure of search still led to successful sensemaking. For example, a sensemaker (MB5) concluded that Trident did not put enough effort toward attracting the youth market after a thorough search that failed to find any advertisements that target the youth. In another example, a sensemaker (MB6) failed to find much useful information about a particular ethnic group as a potential market of a wireless company, so she decided to eliminate that group as its potential market because it would be difficult to do any analysis on that group with no data. In these cases, not finding materials helped users to understand the scope of data that was available and thus helped them to reduce uncertainty.

1b. Successful search, sensemaking failure. Participants failed to make sense of the search results and gave up. Failure of sensemaking may be caused by too much information, too little information, or failure to connect with existing knowledge, in which case the sensemaker were not ready to consume the information found. For example, a user (MJ14) found a seemingly useful article about sustainable energy, but she could not make sense out of it because "*I'm confused, and I don't know if this has anything to do with my story.*" Search results may still be put into note space, but in most cases they were never looked at. In other cases, sensemakers went back to the search results later and were able to make some use of them with their updated knowledge.

2b. Empty search, failed sensemaking. Users did not find any useful information and were not able to make any sense out of the absence of information.

4.2.3.3 Failure Starting Points

The notion of a failure entry point emerged from several cases. Sometimes users started an iteration with an entry point which would not get them anywhere close to their goals.

An example of a failed entry point was when User MB1 started his business analysis case, and immediately started reading the case material from the very beginning, going directly into the details. Soon he realized the details he read did not connect with him. The failure of this entry point occurred because:

1) The sensemaker did not have the procedural knowledge (or knowledge how) of how he would proceed with the case.
2) The sensemaker did not have the necessary initial structure to fit the details that he read in the case.

After the failure of this round's sensemaking, the user analyzed the task, decided how he was going to proceed (primarily bottom-up approach), reviewed some training materials, and created a note page titled "problem of case". He then started reading the "introduction/chronology" session of the case material, and instantiated the "problem of case" concept with several facts in the form of a bullet list. At the end of the iteration, he introduced two more concepts (stock and retirement money of employees) and a causal relationship between them.

In the second iteration, the sensemaker selected an entry point that was more likely to be successful by analyzing the task and choosing an approach (bottom-up). He was clear about the goal of abstracting structures from the new material.

In another example, a sensemaker (MJ6) conducted a key word search “cellular phones” AND “regional” AND “national” in a marketing research database, changed key words several times but did not find anything, and moved on to a Google Web search “regional vs. national cellular company”, but still did not find anything.

This iteration started with an entry point which was likely to fail because the choice of search term “cellular” is not a common expression that would be widely used in Web resources. Moreover, the search was not closely related to the goal of the task, which was to compare several ethnic groups as potential markets to enter.

The user then decided to move on, specified several aspects of the potential markets in CMap, and was able to find information on these aspects.

4.2.3.4 Anticipation vs. Surprise

Some iterations started and ended within the sensemakers’ anticipation, while others started or ended as surprises. More often, sensemakers moved from one iteration to the next by having an idea about what they were going to do next. However, sometimes a gap may have been identified unexpectedly during the current iteration, causing the sensemaker to change the original route of sensemaking and start an iteration working on the new gap.

The iterations that users anticipated had neat paths where they were ended once a part of the sensemaking task was done or declared cancellation. However, when users started a iteration with surprise, they often left the previous iteration abruptly. Some users claimed that they would *"come back to it"* later, but then did not remember to finish what was left behind.

Chapter 5: Conceptual Changes to Knowledge Structure

Conceptual changes describe changes in the sensemaker's conceptual framework (knowledge schemata represented as notes, maps, and texts) that were constructed during the sensemaking process. They constitute an important aspect of the sensemaking phenomenon. Different types of change may occur, including adding facts, modifying existing categories and concepts, and introducing new concepts or relations. The degree of change ranges from gradual to radical changes. Prior knowledge plays an important role in sensemaking. It is the object upon which the conceptual changes occur, and its status or condition influences the sensemaking approach and the sensemaker's ability to make conceptual changes.

This chapter is organized as follows:

Section 5.1 discusses the *types* of conceptual changes, including accretion, tuning, and restructuring of a knowledge schema.

Section 5.2 discusses the *role of prior knowledge* in sensemaking.

5.1 Types of Conceptual Changes

As the sensemakers in the study proceeded with their tasks, their conceptual spaces experienced a series of changes that led to the final representations. This section presents qualitatively different types of changes. Accretion, tuning, and restructuring seemed to be increasingly significant. However, they did not necessarily reflect the degree of significance to the sensemakers' conceptual spaces. For example, adding a less important new concept (restructuring) may not have been as significant to the task as changing a core concept (tuning).

The conceptual space is composed of the internal representations (as evidenced by think-aloud protocols) and external representations (search queries, documents viewed, notes, concept maps, and written reports). The internal and external representations functioned together to provide the working space for sensemaking. The conceptual changes reported in this section include changes to both internal and external representations. See Section 5.3 for discussion of the relationship between internal and external representation.

Changes to external representations were easily observable from the recording of screen activities. I relied on think-aloud protocols to assess sensemakers' internal representations. This section describes the different types of changes happening to both representations.

5.1.1 Accretion: Factual Addition

Accretion refers to the factual increments to the conceptual space. It does not involve structural change. Accretion presumably occurs through appropriate exposure to the information to be acquired (Rumelhart and Norman 1981). As observed in this study, factual addition often took place in different forms of notes and annotations. Accretion constituted the majority of conceptual changes observed in the cases. In this section "fact" is used roughly as "data" or "a piece of information" as opposed to structure; it does not necessarily imply "truth" or "actual existence".

5.1.1.1 Accretion by Copied Notes

One of the most common ways to add factual information to the conceptual space was by making copied notes from the various sources that a sensemaker found.

Copied notes were made by highlighting a piece of information, right-clicking on it, selecting "copy" from the menu, and then pasting the piece of information in the desired place in the notes. Some users used the hot keys "Ctrl + C" and "Ctrl + V". Some used the "drag-and-drop" function to drag a selected piece and then drop it onto the note space. In many of the cases, accretion by copied notes kept the original wording and sometimes the original format of the information. The facts were kept in the original format for two reasons:

1. They were purposefully saved in the original format for "quotation" use later on in the writing stage. This was particularly true for the journalism cases, but was also seen in the business cases.
2. Sometimes the sensemakers' copied these "raw" facts into their conceptual space with the consideration that the facts were somewhat relevant to the sensemaking task at hand. However, they did not have the time or resource, or their established knowledge schemata were not quite ready, for immediate processing. They saved these notes in relatively loosely defined categories (note pages), such as "background and issues" or "problems", and came back to them later. As it turned out some of the facts were never processed or used.

With this type of accretion, the ability to track back to the source of the fact is important. Once sensemakers copied a segmented piece of data, the context in which the data is grounded is lost. Sensemakers sometimes needed to go back to check the source to remind themselves about the context within which the extracted fact appeared.

5.1.1.2 Accretion by Restated Notes

Some sensemakers added factual information to their conceptual space by typing their own notes or annotations to restate what they learned about an object from a source article. Such restatements were mostly present in the following two conditions:

1. The source documents were not in a condition to be copied. Sometimes relevant materials were printed documents. Sometimes the owner of a webpage or document might have disabled the ability to copy-and-paste from a document. Thus sensemakers needed to regenerate the statement of the fact in their note space.
2. The source documents were accessible, but the original wording of the fact did not meet the functional requirements of sensemaking. In some cases, the facts may have reflected a perspective that was not exactly what the sensemaker needed. For example, a fact reporting the lack of market share in certain market divisions in a business annual report could have been restated as an opportunity for a marketing campaign. In other cases, the facts might have appeared in language that was purposefully made vague or misleading, which is often the case in mass media. The sensemaker (journalist) may have needed to read through the surface level message and restate it to reflect her point of view.

In these circumstances, instead of copying the statement of a fact word-by-word, sensemakers used their own language to express the meaning. The new expression of the facts was often made to carry some functional demands of the sensemaking task. The facts were not merely put into a set of unsorted or hardly sorted notes; they were reformed in the process of restatement and often were organized in certain ways in which the functional demands put upon them could be more easily responded to. For example, accretions by restatements were often put onto certain positions such as the top or far right of a note page so that they could be easily accessed later. They also tended to be formatted more than plain copied notes and sometimes put into note groups to carry some functional demands of the sensemaking task.

See section 6.1.1.2 for more details about restatement as a cognitive mechanism.

5.1.1.3 Accretion by Summary Notes

Accretion by summary notes is similar to accretion by restatement in many ways. The facts were reformulated by sensemakers with some consideration about their functions to the sensemaking task. Summary notes were often added when the original articulation of the facts were lengthy. Accretion by summary notes was shorter in length but presumably covered equivalent content of the original.

Sometimes, accretion by summary notes led to discoveries of new concepts or relationships. Summary notes enabled more abstraction than copied notes and restated notes. Such accretions were often made in the form of a bullet list or other similar formats and put in more accessible places in the notes.

5.1.1.4 Accretion in Concept Maps

Factual accretion in concept maps was rare due to limited functionality of the concept mapping software. Often concept maps were created when a general understanding was established such that the facts were already stored in notes during the information gathering stage.

5.1.1.5 Accretion to Internal Representations

Some facts were in the sensemakers' internal space without ever being present in the external representations. Some sensemakers processed a large amount of information in their head and put down in the external conceptual space only the mere amount of information that they needed for the task. For example, when reading a news article, a sensemaker noticed some facts about certain polling results and possibly remembered them, but the facts were not important or interesting enough to put down in the notes for his news article to be written. These facts were added to and processed in the internal conceptual space, but did not leave any trace in the external representations. In some cases, the internal representation could have been considerably larger than the external representation. See section 5.3 for a discussion on the relationships between internal and external representations.

5.1.2 Tuning: Changes to Existing Structure

A sensemaker's understanding of a concept or category is ever-evolving during the sensemaking process as he gradually acquires new information on the concept or category. Tuning involves minor modifications to the existing knowledge structure sensemakers use for interpreting new information. The sensemakers continually tune the knowledge structure to meet the functional demands placed on them. This type of conceptual change involves adaptation of existing structure as opposed to creation of new structure (concepts or relationships).

5.1.2.1 Changing a concept or category's scope

The most common type of tuning is changing the scope of a concept. As a sensemaker accumulates more factual information and learns more about a concept or category, she may find that the concept's original scope does not account for the new information, and she has to adjust the concept's scope to a larger, smaller, or different scope.

Expanding a concept / category's scope. In some cases, the original scope of a concept became too restricted to account for the new information discovered. For example, a sensemaker originally created a note page "problem", but later expanded it to also include "opportunity".

Restricting a concept / category's scope. A concept or category may have a larger scope at the beginning; for example, energy *issues* in the presidential election may include a lot of things such as gas prices, offshore drilling, bio fuels, nuclear energy, and so on. But as the sensemaker in the case study learned more about the issues and formed a general idea about the story she was to write, she restricted the scope of *issues* to include only *domestic drilling, nuclear power, and coal plants and coal-to-liquid fuel*, because these were what she wanted to discuss in her story.

Breaking one concept into two or more concepts. In some cases a concept was broken into two concepts at a later stage. This sometimes reflected the sensemaker's discovery of sub-concepts of a broader concept. For example, User

MJ15 replaced the concept “oil” with two new concepts, “gas prices” and “oil dependency”, after she did more research and thinking on the issue.

Merging two concepts. In some cases two closely related concepts were merged into one concept. A sensemaker did this when she decided that the two concepts were difficult to distinguish conceptually by analysis, or she may do so when she discovered that facts she found often mention the two concepts together. For example, User MB5 merged “problem” and “opportunity” into one concept “problem/opportunity” because most of the facts he found talked about the two together since problems often implies new opportunities for improvement.

5.1.2.2 Changing weight of a concept

Different concepts are presumably of different importance to a sensemaking task; thus each concept has a different weight (or perhaps several weights corresponding to different aspects of importance) in the knowledge schema. Sensemakers changed the weight of a concept when they discovered that a concept was more important than they had originally thought. The weights of concepts could be carried in many different formats, such as a **bold**, colored, or highlighted header or phrase in the notes, a larger node box or a center-placed node box in the concept map. In a business analysis case, the sensemaker (MB4) created a business flow chart of an equipment rental company. After she had laid out all the involved parties of the business, including founder, management, drivers, customers, and so on, the sensemaker figured that “drivers” should get a bigger node to indicate the importance of this concept to the business, because *“the driver has a lot of responsibility… they have to get everything to the right place at the right *

time.” The same process occurred with the concept “customers” because *“they (the company) focus on service”*.

5.1.2.3 Changing the relationship between two concepts

Perhaps due to the limitations of the concept mapping software and the scope or complexity of the concept maps, not much tuning or minor modification to relationships occurred. In principle, the same types of changes should apply, including changing a relationship’s scope or weights.

5.1.3 Restructuring: Creation of new structures

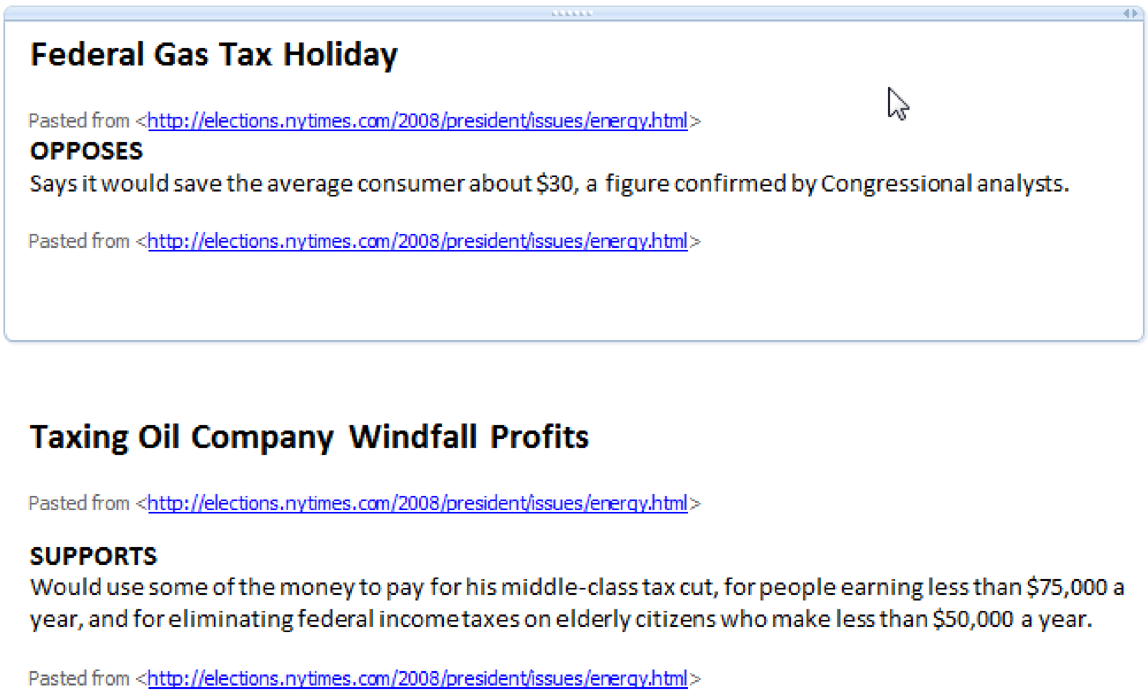
5.1.3.1 Creating a Note Page

Creating a new note page often reflected the recognition of a new concept or category. These concepts and categories formed the top-level structure of the task. They were often broad enough to among them cover most of the relevant aspects of the sensemaking task. They were also general enough to include as much detailed information as was needed in each page.

The structure created in this fashion was usually taken from task analysis or prior knowledge. The sensemakers needed to discover sub-concepts or sub-categories emerging from data to form more useful structures for their sensemaking. Sometimes, new concepts the sensemaker identified from the information found were sometimes important enough to be added to the sensemaker’s top level structure.

5.1.3.2 Creating Structures inside a Note Page

Sensemakers created structures inside a note page by giving a heading for a note box, or a group of note boxes (explicit). This does not require further explanation. For example, User MJ3 used headings for note boxes to separate different energy issues. The following figure shows the structure inside her note page:



Federal Gas Tax Holiday

Pasted from <http://elections.nytimes.com/2008/president/issues/energy.html>

OPPOSES

Says it would save the average consumer about $30, a figure confirmed by Congressional analysts.

Pasted from <http://elections.nytimes.com/2008/president/issues/energy.html>

Taxing Oil Company Windfall Profits

Pasted from <http://elections.nytimes.com/2008/president/issues/energy.html>

SUPPORTS

Would use some of the money to pay for his middle-class tax cut, for people earning less than $75,000 a year, and for eliminating federal income taxes on elderly citizens who make less than $50,000 a year.

Pasted from <http://elections.nytimes.com/2008/president/issues/energy.html>

Figure 5-1: Example Note Page of User MJ3

Structures were also created by grouping notes together into logical units (implicit). This was done in several ways:

- Drawing boxes around notes or dividing notes into different boxes by topic
- Typing in extracted bullets to summarize notes in a page
- Using fonts and highlighting to visualize headings and important concepts
- Modifying structural elements in notes

For example, User MB1 used graphic representation to indicate the casual relationship between stock prices and investment diversification options (shown in the following figure).

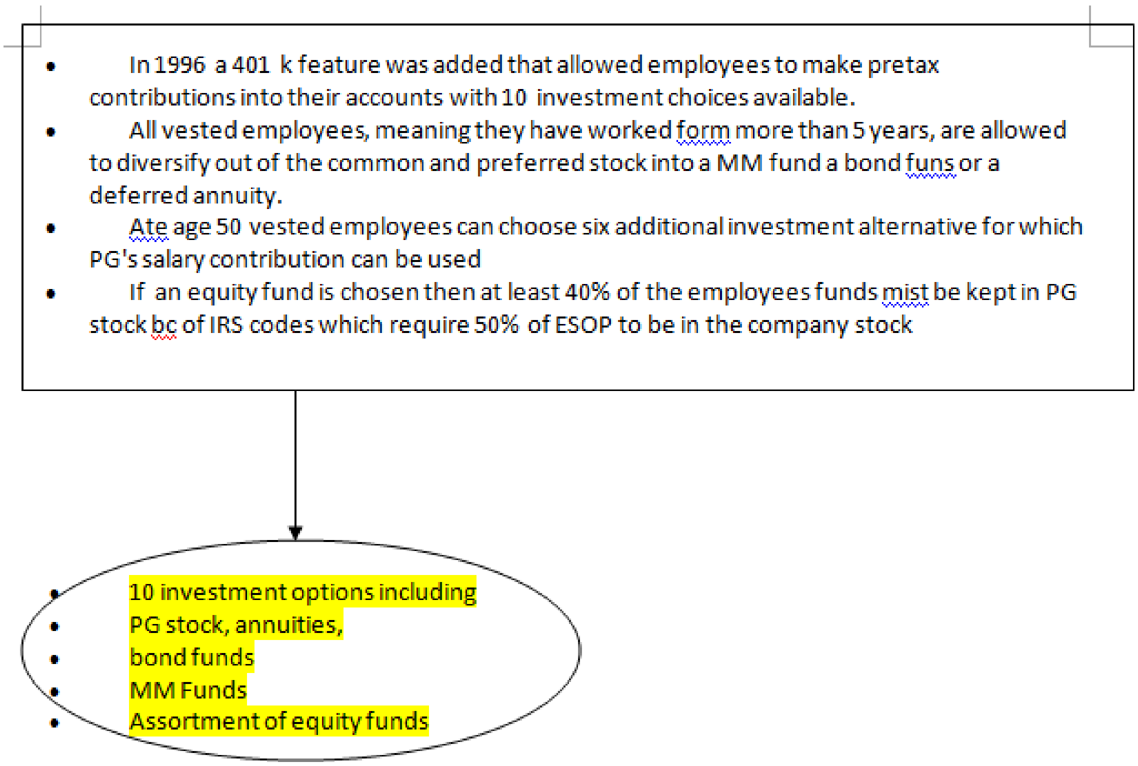


Figure 5-2: Graphic Arrangement of Notes (User MB1)

In another example, User MB5 typed in the following bullet list to summarize the note on the "Problem/Opportunity" page:

- The bolder the flavor the better
- Older audiences wants sugarless, youth not as much
- Older people want oral care gums
- Caffeinated gum for younger groups

Figure 5-3: Summary Notes (User MB5)

5.1.3.3 Creating new concepts and relationships in CMap

New concepts were identified when sensemakers created concept maps to visualize what they had learned so far. The process of mentally mapping the concepts and laying out their relationships required further analysis that sometimes resulted in the creation of new concepts. For example, User MJ3's concept map included two nodes "presidential debates" and "issue of economy defeating the environment", which were not part of the structure represented in her notes.

5.1.4 Changes in Representation

Changes in representation happened when sensemakers simply transformed representation in one form (for example, notes) to another (for example, concept maps). Changes in representation altered only the form of the structures, but did not affect the fundamental knowledge structure. Changes in representation were partially due to the limitations of the sensemaking tools not being integrated. The sensemakers had to manually recreate a map based on their notes. However, the process of changing representation from one form to another actually helped some sensemakers to recognize gaps in knowledge, discover new concepts, and form new relationships. Figure 5-4 illustrates the note structure and concept map created by User MJ3.

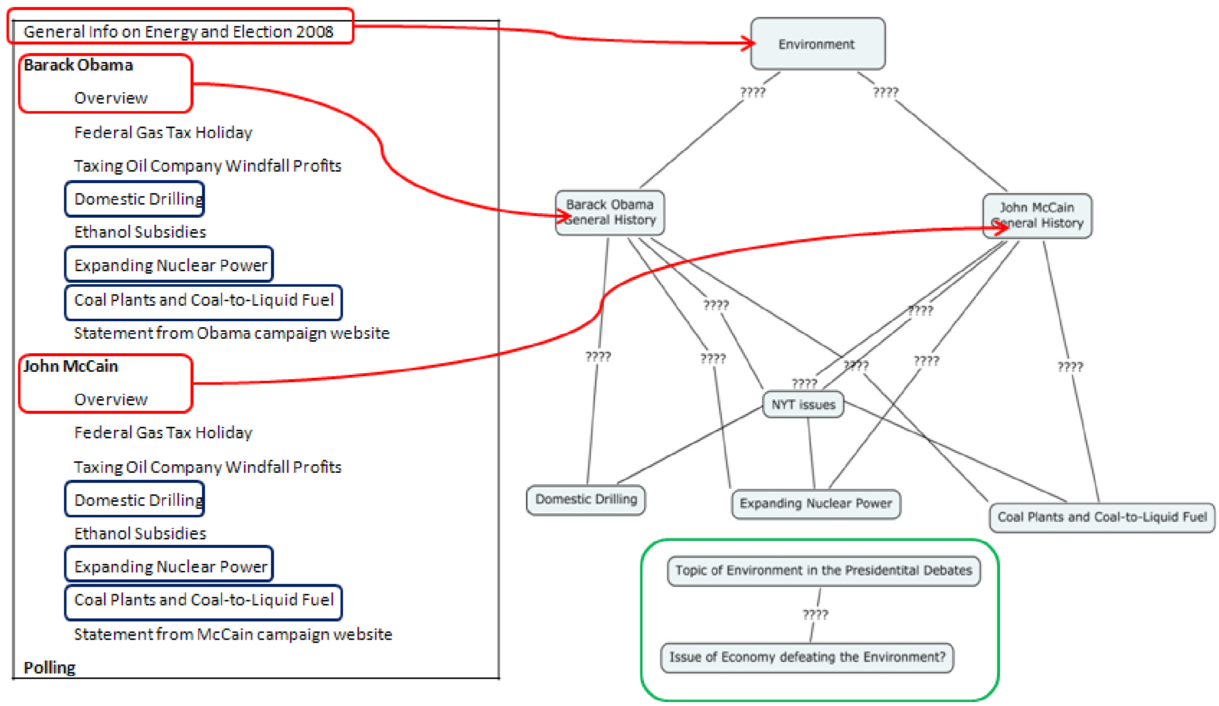


Figure 5-4: From Notes to Concept Map (User MJ3)

Regardless of the specific formats, the most commonly observed changes in representation include:

1. One-to-one mapping: a concept or relationship was mapped to a concept or relationship in the other representation format;
2. Many-to-one mapping: multiple concepts or relationships were mapped to a single concept or relationship in the other representation format;
3. Selective inclusion: among a group of concepts or relationships in the original representation format, only a few concepts and relationships were selectively included in the other representation format;
4. Making implicit concepts and relationships explicit.

5.1.4.1 One-to-one Mapping

Most users constructed their concept maps after they have collected information about created note pages to organize the information they collected. The concept maps were not constructed from scratch but were transformed from the structure and data represented in sensemakers' notes.

When constructing a concept map from notes, the top level structure (often represented as note pages) was often kept intact for the most part. A one-to-one mapping was often observed among the top-level concepts represented in the notes and in the concept map. Sometimes the concepts were labeled slightly differently in the new representation format. In the example shown in Figure 5-4, 4 out of a total 12 concepts in the note structure had a one-to-one mapping in the concept map. Table 5-1 summarizes the one-to-one mappings.

Table 5-1: One-to-One Mapping of Concepts (User MJ3)

Concept in Notes	**Concept in Concept Map**
General info on Energy and Election 2008	Environment <a pars pro toto label>
Domestic Drilling	Domestic Drilling
Expanding Nuclear Power	Expanding Nuclear Power
Coal Plants and Coal-to-liquid Fuel	Coal Plants and Coal-to-liquid Fuel

In another example, User MJ9's concept map and notes also demonstrated one-to-one mapping of some concepts. Table 5-2 summarizes this mapping:

Table 5-2: One-to-One Mapping of Concepts (User MJ9)

Concept in Notes	Concept in Concept Map
Background and issues	Energy
Obama	Obama
McCain	McCain
Polling Results	Public Opinion

In both examples, the concepts were higher level (top and second level) concepts. Sometimes the concepts were labeled slightly differently in the new representation format. In the second example, "background and issues" were labeled "energy" but the think-aloud protocol suggested that the concept was intended to capture the background and issues on the energy topic. "Polling results" were labeled "public opinion" in the concept map which was later instantiated with polling results.

5.1.4.2 Many-to-one Mapping

In some cases, instead of having a one-on-one mapping for each concept, multiple concepts in Notes were mapped to a single concept in the concept map, or vice versa. Table 5-3 shows two examples of many-to-one mapping from two concepts in Notes to a single concept in concept map.

Table 5-3: Many-to-One Mapping of Concepts (User MJ3)

Concept in Notes	**Concept in Concept Map**
Barack Obama	Barack Obama General History
Overview	
John McCain	John McCain General History
Overview	

In this example, a concept at the top level (Barack Obama) and its sub-concept at the second level (overview) were mapped into a single concept "Barack Obama general history". This is an example of a change in representation (transforming notes into a concept map) led to a minor change in structure (tuning).

5.1.4.3 Selective Inclusion

In several cases, when transforming notes into a concept map, users did not include all concepts and relationships appearing in the notes in the concept map. In the example shown in Table 5-4, of all the six energy related issues that User MJ13 took notes on, she only selected three in her concept map.

Table 5-4: Selective Inclusion of Concepts (User MJ3)

Concept in Notes	**Concept in Concept Map**
Federal Gas Tax Holiday	-
Taxing Oil Company Windfall Profits	-
Domestic Drilling	Domestic Drilling
Ethanol Subsidies	-
Expanding Nuclear Power	Expanding Nuclear Power
Coal Plants and Coal-to-Liquid Fuel	Coal Plants and Coal-to-Liquid Fuel

The selective inclusion was decided based on the importance of the concepts to the task *"I do not think these were important to include in my story"*, *"I do not think I can talk about all of them"* (User MJ3). The important concepts included in the concept map were also highlighted in her notes.

5.1.4.4 Making Implicit Concepts and Relationships Explicit

When transforming from notes to concept maps, sensemakers' sometimes made implicit concepts and relationships explicit by creating a node or link in the concept map. In many cases, these implicit concepts and relationship were hidden in the text format in notes. For example, Figure 5-5 shows User MJ9's note structure. It includes only 6 top-level concepts (as note pages). There is no explicit structure (such as headings and sub-headings or note boxes) within each note page.

Background and issues
Obama
McCain
Obama Quotes
McCain Quotes
Polling Results

Figure 5-5: Note Structure created by User MJ9

When transforming the notes into a concept map, for the notes on “McCain” and “Obama”, User MJ9 defined two relationships: “for” and “against”. She also made explicit the various energy related issues such as “offshore drilling”, “nuclear energy”, “reduced gas tax”, “oil company tax”, “ANWAR drilling”, and “Strategic Pet. Reserve Release”. She linked “McCain” and “Obama” with these issues using the relationships “for” and “against”. Figure 5-6 (next page) shows the concept map created by User MJ9.

These concepts and relationships did exist in the notes in a text format. For example, a paragraph under note page titled “Obama” reads *“Sen. Barack Obama stuck to his planned script today, pledging, if elected president, an aggressive effort to reduce American dependence on foreign oil by investing in research on alternative fuels and relying more on nuclear, wind and other energy sources.”* This text representation was transformed into a relationship “Obama <is for> nuclear energy”. The main concepts and relationships were made more salient through the transformation.

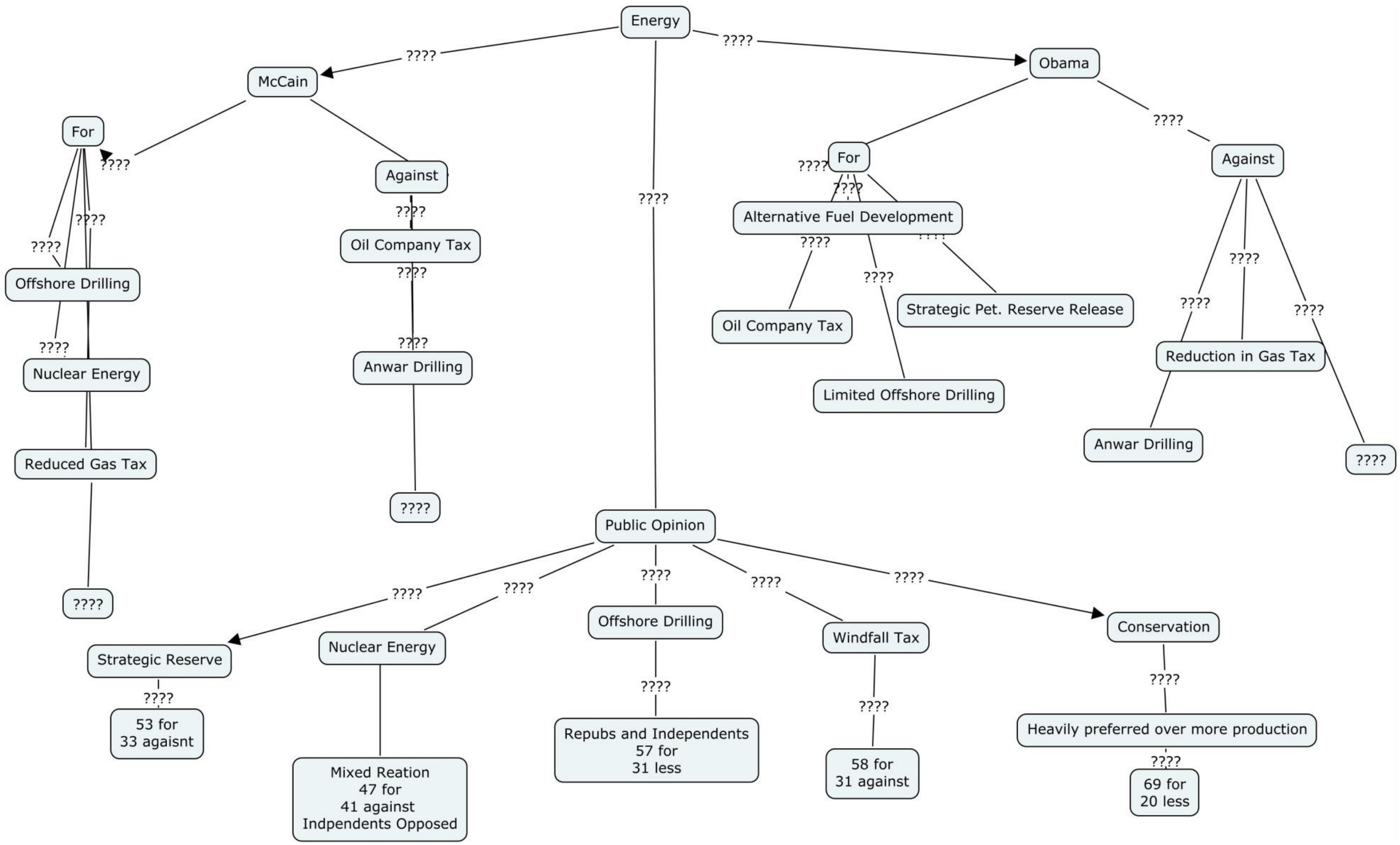


Figure 5-6: Concept Map created by User MJ9

5.1.5 Relationships among Different Types of Conceptual Changes

The main difference between tuning and restructuring is how fundamental the changes are. Accumulation of factual addition and tuning does not necessarily result in restructuring, although failure of tuning to account for new fact might trigger restructuring. For example, when finding an article about gas prices from an online news press, User MJ6 found none of the existing note pages he created, “background on energy”, “Obama”, “McCain”, and “Polling Results”, seemed to be a good place to put the article. He tried to redefine the “background” note page and put the article under “background” but the article about gas prices does not seem to fit with the other articles under that page which are more general introductions on the issue of energy. As a result, he created a new note page titled “gas price” as a special energy related issue and instantiated the note page with the article.

Changes in representation may result in tuning or restructuring. For example, when creating a concept map from notes, User MJ3 merged top-level concept “Obama” and second level concept under the “Obama” note page “overview” into one single concept in the concept map “Obama general history”. The changes in representation resulted in tuning. For another example, User MJ3’s concept map included two nodes “presidential debates” and “issue of economy defeating the environment”, which were not part of the structure represented in her notes. These ideas were expressed in the articles she copied into her notes but was not represented as part of the structure. This is an example of changes in representation resulted in restructuring.

With tasks that are complex enough, sensemakers may need to conduct multiple tuning steps accounting for new data, and the multiple tuning steps may result in restructuring. The cases in this study were not complex enough to result in multiple tuning or restructuring steps to account for any piece of new data.

5.2 Prior Knowledge

The acquisition of new knowledge makes little sense without assuming some prior knowledge within which new information is interpreted; otherwise the new knowledge will be unintelligible (Rumelhart and Norman 1981). Prior knowledge was the base upon which all the conceptual changes took place. It also seemed to influence and direct sensemakers' approaches.

5.2.1 Relationship of Prior Knowledge to what needs to be Known

Prior knowledge plays an important role in learning (Rumelhart and Norman 1981; Wittrock 1990; Anderson, Reder et al. 1996; Grabowski 1996). A person may have prior knowledge about the content to be learned, or he may have the pre-requisite knowledge to understand what will be learned. In sensemaking, the distinction between the pre-requisite and to-be-learned knowledge is not as obvious. For example, for a sensemaker to be able to make sense of information found on energy and election issues and write a story on that, he needs to have some general knowledge about how election works, why energy and other issues are important in presidential elections, and how to write news articles. The boundary of what is pre-requisite and what is to be learned is not as clear as in science learning.

For science learning, Chi (2007) identified three conditions of prior knowledge in relation to the to-be-learned knowledge: missing, incomplete, and in-conflict-with. This study adopted these conditions but defined them in relation to the needed knowledge for the sensemaking task:

- Missing: the sensemaker has no knowledge about a concept in the task domain.
- Incomplete: the sensemaker has some knowledge about a concept in the task domain, but additional search and sensemaking are required to accomplish the task.
- In-conflict-with: the sensemaker has some knowledge about a concept in the task domain, but his knowledge is inconsistent with what is needed to accomplish the task.

Figure 5-7 illustrates these conditions in regards to Figure 4-1: Gap Identification, for the simplified case where the user is aware of all she knows and she needs to know.

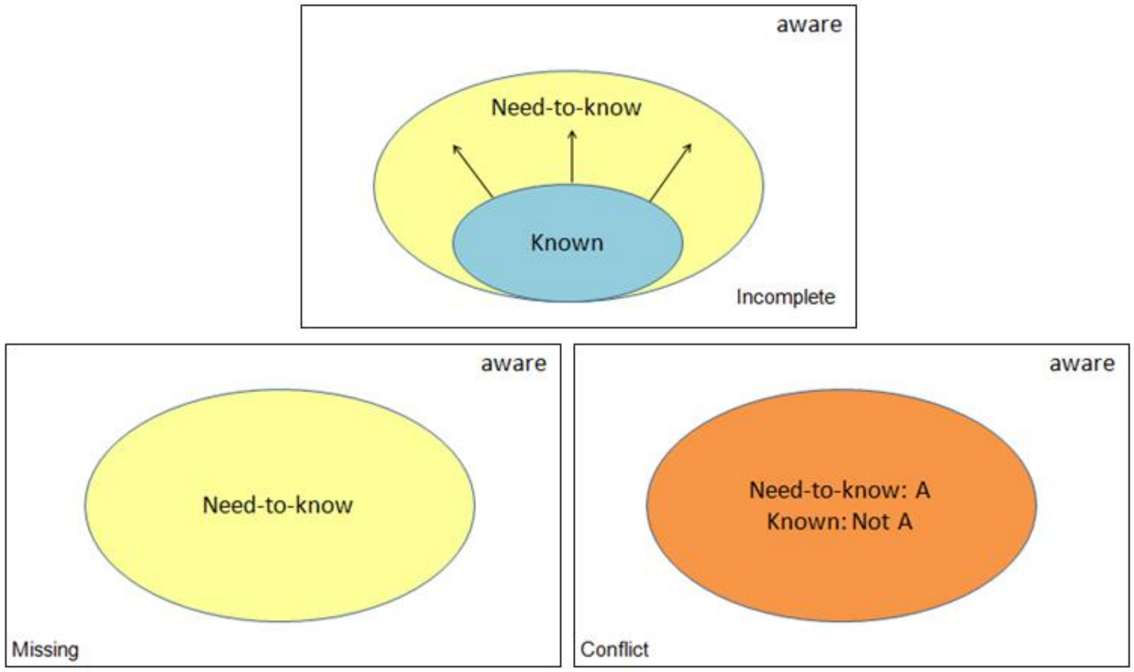


Figure 5-7: Three Knowledge Conditions

The top box shows the "incomplete" condition, which was the most common. The sensemaker had some knowledge about what they needed to know, but their knowledge was incomplete. For example, a sensemaker (MJ4) "*knew a little bit, because one of my summer classes I did sort of a general paper with a group about McCain and Obama's' platforms on various issues and one of our issues was energy, so I did recall some but not in detail, so despite of my fogginess I was able to look a little more detail into the issue"*. In fact several users with the news writing article mentioned that they "*knew a fair amount about the energy policy in general"* but were *"not exactly sure where the candidates stood."*

The bottom-left box shows the "missing" condition. The sensemaker did not know anything about what they needed to know. This was rarely the case in this study. Although in a very few cases, users did not have much specific knowledge needed for the task except for some common-sense knowledge about the domain. For example, when asked about what User MB5 already knew about the topic (gum marketing), he said *"not much, really. Just general knowledge about what I have seen, gum and everything, but not really much at all."* It is not accurate to claim this user's prior knowledge was in the "missing" condition, because general knowledge about gum is also needed for the task. The "missing" condition may be very rare in everyday life sensemaking scenarios too. Usually sensemakers start a sensemaking task knowing something about the topic domain; otherwise they will not have the initiative to start a sensemaking task. Sensemakers with learning tasks that are assigned to them may know very little about the subject domain and are probably as close as one can get to the "missing" condition.

The bottom-right box shows the "conflict" condition. The sensemaker had prior knowledge about the topic, but his knowledge was incorrect, or "in conflict with" what needed to be known. For example, a sensemaker (MJ9) started the task thinking that people were against offshore drilling and surprisingly found out that "*a lot of people seem to be really favoring drilling.*" Similar to the "missing" condition, it was very rare that a sensemaker held complete misbelieves and had to correct everything that she knew about a topic.

More often, users knew something about the topic that they worked on, though a small part of their prior knowledge was in conflict with what needed to be known. Taking the conflicts into consideration, a modified figure for gap-identification is shown in Figure 5-8.

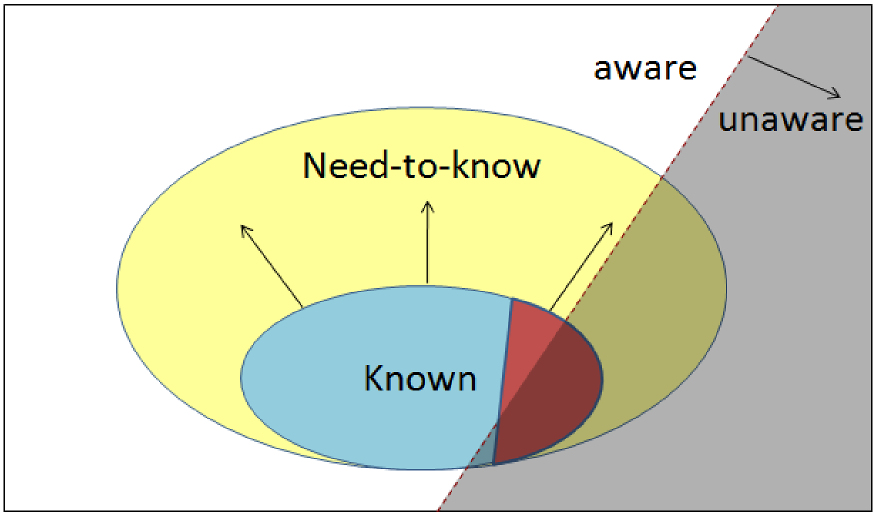


Figure 5-8: Gap Identification, Considering Conflicts

When expanding prior knowledge to what the sensemaker needs to know (possibly after or while increasing awareness of what she needs to know), the sensemaker should be sensitive to the fact that in some cases the needed knowledge required may be in conflict with her existing knowledge and correct her "incorrect" existing knowledge.

In addition to expanding prior knowledge to what is needed to know and increasing the awareness, the sensemaker should also identify whether there is any conflicts in her prior knowledge to what she needs to know and correct the "incorrect" knowledge. Often this can be done only by seeing new information.

5.2.2 Relationships of New Information to Prior Knowledge

Researchers have long recognized the importance of prior knowledge in learning. Most of the learning that occurs in life is either incorporated within prior knowledge (Piaget's assimilation) or modifies prior knowledge (Piaget's accommodation). The user study discovered four relationships of new information to users' prior knowledge.

a. Agreement: the new information is in agreement with prior knowledge. This type of information is easily accepted and will reinforce prior knowledge. For example, "*so again, just 'older people want oral care gums', OK, just the same stuff over and over again.*"(User MB5)

b. Complementary: the new information completes prior knowledge. If prior knowledge is incomplete, what is mostly needed is complementary information. "*This is good stuff! I did not know this. 'Could claim a tax credit of up to 10 percent of the cost of all qualified...'*" (User MJ14)

c. In-conflict-with: the new information puts prior knowledge in doubt. This is not going to help condition 1 and 2; but large amount of this type of new information may be able to correct misconceptions or wrong beliefs in condition 3. "*He (McCain) is all about fuel efficient vehicles. Which, in the past, you know, if you buy into the stereotypes of Republicans, they drive the big cars, right? Well John McCain says, 'No, not right, we drive fuel-efficient cars!' So let's see if something actually gets done about that. It will be very interesting. I'm obviously very cynical towards both sides of the government.*" (User MJ13)

The following table summarizes how useful the new information is to sensemaking in different conditions of prior knowledge.

Table 5-5: Usefulness of New Information in Different Knowledge Conditions

Condition of prior knowledge in relation to needed knowledge	Relationship between new information and prior knowledge		
	a. Agreement	b. Complementary	c. In-conflict-with
1. Missing	N/A	**Most useful**	N/A
2. Incomplete	Could be useful	**Most useful**	Could be useful
3. In-conflict-with	NOT useful	Could be useful	**Most useful**

When needed knowledge is missing, new information that connects to what the user knew may help the user to recognize and find what is missing. If the new information is connected to a sensemaker's prior knowledge, it is more likely to be recognized and incorporated into existing knowledge.

When prior knowledge is incomplete, new information that complements the prior knowledge is most useful.

When prior knowledge is in conflict with what is needed to know, information that could potentially correct the false believes is most useful.

5.2.3 Dealing with Conflicts

When the information was in agreement or fit with the participants' existing knowledge, sensemakers often did not mention that in their think-aloud protocols; when conflicts happened, participants often explicitly talked about how they dealt with the conflicts and the results of that conflict. This observation indicates that conflicts can pose serious challenges to sensemakers because conflicts do not conform to what they knew. When participants detected conflicts, either between two pieces of information they found or between new information and their existing knowledge, there were four outcomes out of how the conflicts were dealt with (Plous 1993):

1. *Disregard*: The participant refused to accept conflicting evidence and kept the original conceptual representation; no conceptual changes happened. *"… I wanted that article to say something else. I have to disregard it."*
2. *Compromise*: The participant partially accepted the conflicting evidence and partially changed his or her existing conceptual model to integrate the new evidence. User MJ13 surprisingly found that "*A lot of people seem to be really favoring drilling"*, although contrary to her original belief, she compromised and adjusted her understanding *"I can only assume that would come from the desire to note be dependent on foreign oil… Which everyone wants kind of across party lines."*
3. *Acceptance*: The participant updated his/her existing representation and accepted the new evidence. As a result, the existing conceptual model often had to be restructured. "*Oh here we go 'concern over imported oil from the Middle East', and the newfound muscle of California's eco-voters. Oh, okay, that's good. That's good to know.* " (User MJ14)
4. *Confusion*: The participant failed to resolve the conflicts. *"…Obviously I have no idea what this is about..."* This often resulted in an unstable or unsatisfied mental state. Sometimes the participant searched for more information until the confusion was transformed by

new evidence into one of the above three states. Sometimes a participant simply gave up and moved on.

The ways users dealt with conflicts might be influenced by the strength and coherence of the new information and the strength, coherence, and degree of commitment to the users' previous knowledge (Dole & Sinatra, 1998). How these factors influence the outcomes and what other factors contribute to the acceptance or discard of a piece of information need further investigation.

Chapter 6: Cognitive Mechanisms

This chapter discusses the cognitive mechanisms that users used in their sensemaking. The cognitive mechanisms serve a variety of purposes and enable sensemakers to tackle the problem of sensemaking by different means. The seventeen unique cognitive mechanisms can be divided into four categories by their primary functional use:

1. Mechanisms used in processing new material
2. Mechanisms used in examining individual concepts
3. Mechanisms used in examining relationships (relationships among concepts and relationships among facts)
4. Mechanisms used in examining anomaly and inconsistency

Some mechanisms are mostly data-driven, such as mechanisms used in processing new materials like key item extraction, restatement, and summarization. Some mechanisms are mostly logic driven, such as inference, explanation-based mechanisms, and specification. Others are somewhat in-between: some may require both data and logic to function, such as classification and semantic fit; others may be used upon either data or structure, such as comparison. However, these categories are not completely isolated from each other; even the data-driven mechanisms may be influenced by structure.

The following table shows the cognitive mechanisms in their functional use and data-driven or logic-driven quality:

Table 6-1: Cognitive Mechanisms

	Data-driven	**Logic-driven**	**Neither**
Processing new information	• Key item extraction • Restatement • Judgment evaluation • Summarization • Generalization • Schema induction	• Schema induction	
Examining concepts		• Definition • Specification	
Examining relationships	• Comparison • Analogy/Metaphor • Stereotyping • Classification	• Comparison • Analogy/Metaphor • Stereotyping • Classification • Explanation-based • Inference	
Examining abnormality and inconsistency	• Semantic fit	• Semantic fit • Elimination	• Socratic dialogue

The cognitive mechanisms are of different complexity. Some mechanisms such as key item extraction in most cases require little extra effort from the sensemaker. Some are more complex, and the sensemaker must purposefully concentrate on them. One way to measure complexity is by the number of items that a sensemaker has to deal with at the same time. For example, comparison requires that the sensemaker deal with at least two items at the same time, whereas restatement deals with only one item. Some mechanisms deal with multiple items. For example, semantic fit requires the sensemaker to examine multiple items at the same time to determine the fitness of a concept and its relationship with other concepts in the knowledge structure.

The following figure illustrates cognitive mechanisms that range from data-driven to logic-driven (as illustrated by the X-axis), and across differing levels of complexity, measured by the number of items dealt with at one time (Y-axis).

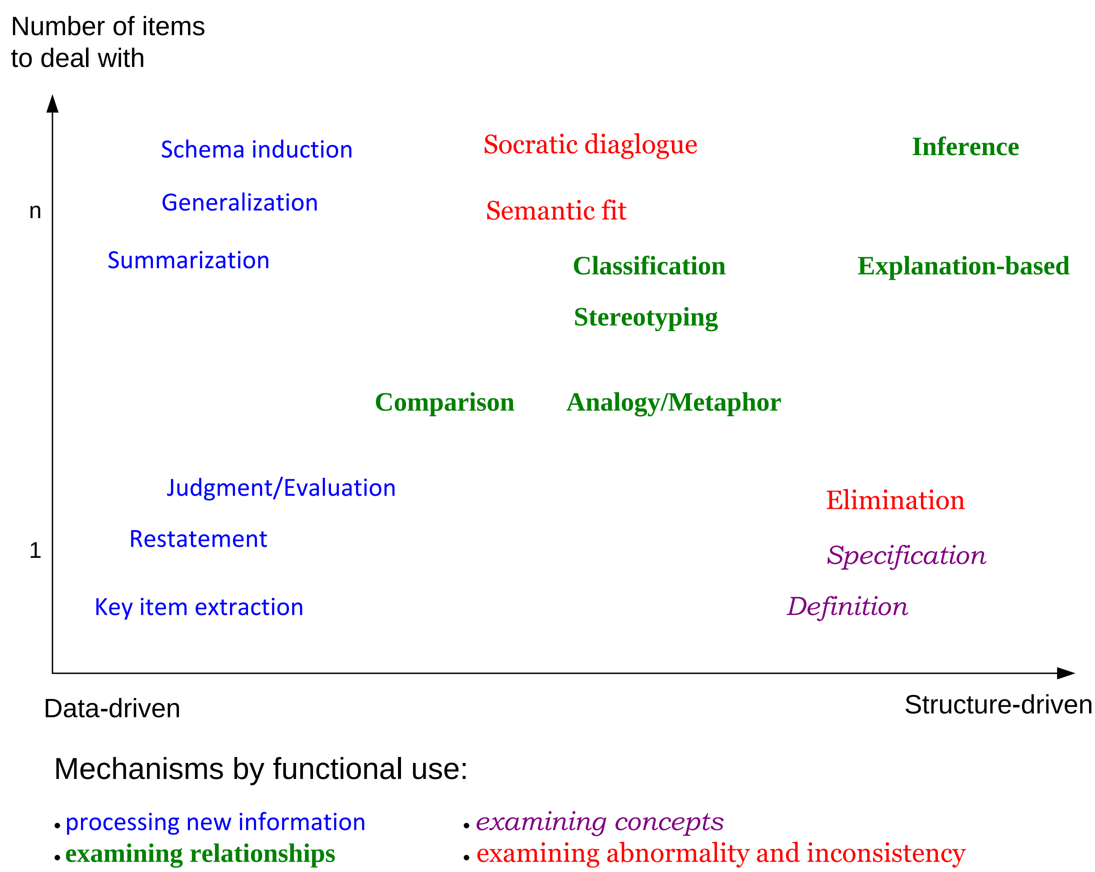


Figure 6-1: Cognitive Mechanisms Used in Sensemaking Tasks

The mechanism may be used alone or in combination to move the processes along and to trigger conceptual changes. The degree to which a sensemaker uses a bottom-up or top-down approach can be analyzed at two levels: task level and the step level. In most cases, sensemakers did not commit to just one approach, but used a combination mode instead. Many factors may influence the approaches that a sensemaker uses, such as characteristics of the tasks, cognitive styles of the user, and the user's prior knowledge.

Section 6.1 talks about the cognitive mechanisms by their functionality and their uses. Section 6.2 discusses the mechanisms from the data-driven vs. logic-driven perspective. Section 6.3 discusses users' sensemaking approaches (bottom-up and top-down).

6.1 Mechanisms by Functionality

This section discusses the functional use of cognitive mechanisms. The following table gives an overview of the uses of cognitive mechanisms under each category:

Table 6-2: Functional Use of Cognitive Mechanism

Cases	**All**	**Instances of cognitive mechanisms used in:**			
		processing new information	**examining concepts**	**examining relationships**	**examining anomaly and inconsistency**
MB1-C1	50	37	1	8	4
MB1-C2	10	9	-	1	-
MB3	27	25	-	-	2
MB4	62	45	4	9	4
MB5-C1	127	67	7	39	14
MB5-C2	89	62	3	18	6
MB6	32	25	2	2	3
MJ1	34	30	-	1	3
MJ3	67	50	2	7	8
MJ4	13	6	-	7	-
MJ5	14	10	-	4	-
MJ8	26	21	-	1	4
MJ9	43	24	5	9	5
MJ13	167	124	5	34	4
MJ14	26	21	2	1	2
MJ15	11	6	2	2	1
Average	50	**35**	**2**	**9**	**4**

6.1.1 Mechanisms Used in Processing New Information

Cognitive mechanisms under this category were used mostly when sensemakers encountered new information (both data and structure) and were trying to understand, consume, and build structures from the new materials. The major sensemaking task in processing new material is to build (or extract) structure from the new material, and connect the structure or data to the existing knowledge. The two major purposes that these mechanisms serve include managing complexity of the new material and making it more accessible (mentally) to the sensemaker.

6.1.1.1 Key item extraction

Key item extraction is the most-used mechanism for almost all users. It can be used in extracting concepts and entities; it can be used in extracting relationships. In some cases, it was the sole mechanism used for discovering new structure or facts. The participants extracted different types of key items from the mass of new information they found.

Extracting key (important) words or phrases directly. These key words or phrases included names of people, organizations (such as companies, banks, and news agencies), brands and products, and so on.

Extracting key concepts. Sometimes sensemakers extracted key concepts that were important to the task. For example, participants with the energy news task extracted issues such as offshore drilling or nuclear energy regardless of how they were phrased. Sometimes they may not have used the exact words or phrases as they appeared in an article.

Extracting key relationships. A more complicated extraction was to extract key relationships from a sentence, a paragraph, or an article. For example, User MJ3 extracted Obama and McCain's positions (oppose or support) to the various issues from an article from New York Times, and put the relationships in her notes.

Users used key item extraction alone or in combination with other mechanisms. Sometimes it was the preceding mechanism on which other mechanisms such as summarization and restatement were built.

Key items were extracted at different stages to assist sensemaking. It may have been done by reading aloud key phrases while doing Web search to decide whether some article was worth looking at. It may have been done by highlighting phrases in the browser when reading a relevant article. It may have been done by copy-and-pasting or typing the key items into notes. A few users mentioned typing in important points as an alternative to highlighting in or on the material (especially when the material was printed) "*And then under the heading I was just basically kind of highlighting but instead of highlighting I would write it in, so I could have the important points.*" (MB4)

The boundary between key item extraction and restatement is sometimes blurry when the extracted item is an idea, rather than a specific name. Often when extracting the idea, the sensemaker restates it using her own language. In fact the two mechanisms were often seen used in combination. See section 6.3 for more details.

An item needs to have certain qualities to be extracted by the sensemaker. Any of the following can be considered a *key* item:

- Things the sensemaker has not seen before. New concepts and novel ideas are likely to catch the sensemaker's eye. A few examples of new or novel key items included:

  "… but one thing I think is interesting is his 'Clean Car Challenge.' I had never heard about that." (MJ4)

"Cheap devices..." Don't know what that is. (MB6)

- Concepts important to the task. Focused sensemakers kept in mind what the task required and looked for concepts that would satisfy those requirements.
- Major concepts or entities being discussed in an article. The concepts or entities that were most discussed by an author or had a lot of significance put on them tended to naturally get the attention of the users. Authors of articles or websites might have emphasized these things through use of format (such as bigger fonts or center position), or through language use.

6.1.1.2 Restatement

Sensemakers often restate what was read to put it into a language that is more accessible. With restatements sensemakers reiterate the essential information and ideas expressed by the original but presented in a new form. Restatements applied mostly to statements (or instantiated structures). Compared to summaries, restatements usually did not change the original level of detail. There were two major characteristics of restatements that are relevant to users' sensemaking:

First, restatements may have a different focus than the original, have a new perspective, and/or use a different language. For example, after reading this sentence about Orbit, a competitor of Trident: *"After a couple of decades of success in the UK, sugarless gum brand Orbit launched in the US in 2001 and quickly established a place for itself among the category leaders,"* User MB5 restated this fact as, *"Orbit just has* ***only*** *been in the US since 2001. Wow, that's seen* ***huge growth****, um, (it) was in the UK, that's amazing."* The restatement focused on Orbit's relatively short history and huge growth in US. Furthermore, it was stated in much less formal language than the original.

Second, restatements were often slightly shorter. For example, the same user read the following statement:

"The successful string of new flavor line extensions and new brand launches in the dynamic and highly competitive sugarless gum segment are largely attributed to this market growth."

He restated it as *"That's what I was saying how sugarless gum is the new product in the category."* Again, the new expression is more accessible, with less formal language, and is shorter than the original. It did not pick up "new flavor line" but focused on "sugarless gum", which later became one of the marketing campaigns proposed by the sensemaker.

To summarize, the mental process required for successful restatement with paraphrases not only helps sensemakers to grasp the meaning of the original, but also makes it more accessible to the sensemaker, while lowering its level of complexity with shorter and simpler language.

6.1.1.3 Judgment or Evaluation

Some sensemakers followed their use of key item extraction or restatement with a judgmental or evaluative claim. This allowed them to form opinions toward the information being processed, and connect it with their existing knowledge. The following example is taken from the think-aloud protocol of a user who was working on the election story:

"Short term relief to American families facing pain at the pump," ***which we saw in all those polls, but I'm sure they both agree*** **on that.** *…"Within 10 years save more oil than we currently import from the Middle East and Venezuela combined."* ***That's really lofty.*** (MJ13)

In the above example, the judgments are marked in bold. The first comment indicated that the polling result that the sensemaker was reading conformed with what she had already known about the issue, and suggested that she thought the two candidates had similar positions on that issue. The second comment suggested an unsatisfactory attitude toward the fact, which would lead to either search for more substantial plans or a critical opinion expressed in the article if such plans did not exist.

Sometimes sensemakers made sarcastic comments. For example, while reading the polls, the same user (MJ13) made the following comments (marked in bold):

"Relaxing some environmental standards. Oh, 50-50." <sarcastic> ***Excellent*** *</sarcastic>. "Rationing gasoline and oil." <sarcastic>* ***Opposed because that's ridiculous, obviously.*** *</sarcastic> "Increasing federal tax"* ***Opposed! No one wants***

to pay more money so obviously everyone just wants to blame somebody. They don't want to have to pay for it.

Such comments or judgments helped the sensemaker to form critical opinions. Sometimes they helped the sensemakers to detect certain themes or patterns if similar judgments were made over and over again.

6.1.1.4 Summarization

Summarization was one of the most frequently used mechanisms by almost all users. By summarizing lengthy elaborations, the sensemakers were able to abstract one level up with concise and more easily accessible language. Summarization assisted sensemaking in the following ways:

First, summarization significantly reduced the amount of information to be further processed. For example, after reading several paragraphs about McCain's energy plan, user MJ13 summarized *"Ok, so clearly… he (McCain) wants to break dependency on foreign oil like everyone knows."* The lengthy elaboration was summarized into one sentence and was connected to her existing knowledge (indicated by *"like everyone knows"*).

When creating summaries instead of copying and pasting the original text in their notes, users reduced their workload by avoiding reading a lengthy text chunk a second time. For example, User MB5 suggested that he purposefully limited his notes to one to two sentences so that once he finished collecting information he would not have to read through large chunks of copied notes again.

Second, summarization pulled out main points. For example, after reading a poll about the energy problem, user MJ13 summarized *"Bush administration, oil* *companies, and commodity speculators being the top three reasons for this."* She covered the major points with a condensed version of the original.

Third, summarization included important details. For example, the same user (MJ13) included this detail in her summary: *"Republicans… John McCain supports a cap and trade system to reduce emissions back to 'before 1990' levels. Neat."* The "cap-and-trade" system was one of the important details that distinguish McCain from Obama in this user's view.

Summarization allows sensemakers to reduce complexity in processing a large amount of new information. A sensemaker can manage the scope of sensemaking by getting the main points of an article and leaving out unnecessary details. He can also make the article more accessible by using his own language. For example, a sensemaker can use a less technical language than the author of the original article used.

6.1.1.5 Generalization

Sensemakers sometimes drew general conclusions or claims from a sample, or came to a general conception or principle from particulars. Generalization helped sensemakers to step back from specific facts and move to a higher level of abstraction.

Making general claims from a sample. Depending on the availability of source information, sensemakers sometimes had to accept evidence from a sample to make general claims about the population. This was particularly true for users in the news writing task. They used polling result as a way to get to public opinions of

Americans. Although sometimes they recognized *"this may not be very reliable"*, in most of the cases, they made general claims based on polling results, for example:

*"They're going to have very similar plans when it comes to the environment because **the entire country** – bipartisanly – wants to fix the environment." (MJ13)*

Coming to general conception from particulars. Sensemakers sometimes made generalizations about trends based on particular facts they saw. For example, after seeing facts about the use of gum in adult and youth audiences, a sensemaker (MB5) suggested that *"we're starting to see some **trends** or some **themes** running through all these,"* which he generalized as bold flavors and coolness factor for kids and health factor and whitening for adults. The general conception that youths enjoyed bold flavor and coolness and adults enjoyed the health and whitening benefits of gums helped in shaping the marketing plans into two major campaigns.

Generalization allows sensemakers to move up from the details and focus on the concepts and relationships at a more general level. This helps sensemakers to build structures from data by transforming the specific data into a general claim that focuses on key concepts and relationships.

An example of discovering a general concept from specific data is that after seeing a Trident ad from the TV sitcom "*Friends*", user MB5 created a concept "*TV Ad Placement,*" which is an important aspect of a marketing plan, but the concept was not brought up anywhere else except for the particular placement (i.e., "Friends") that the user found.

For another example, user MB4 generalized from the fact that two people from a company and a bank met and moved the company's account to the bank, to the general idea that *"it is a mutually beneficial relationship for them because they are both getting something out of it."* This generalization allows the user to identify the financial relationship between the company and the bank, which is a major factor in the company's business.

6.1.1.6 Schema Induction

Sensemakers' used schema induction to recognize shared elements between two or more related concepts, phenomena, or situations (Gick and Holyoak 1983). The discovery of the regularities in the co-occurrence of certain phenomena or the recognition of shared elements between concepts, and relationships helps the sensemaker to reach a higher level of abstraction.

Sensemakers induced a schema or frame structure based on what two concepts share in common. For example, when examining different ethnic groups as potential markets for a mobile phone company, user MB6 identified several aspects that she needed to look up, and created a node for each aspect in her concept map, including *"Size", "Growth Rate", "Potential Revenue and Profitability",* and *"Barriers to Entry".* By examining each ethnic group using these aspects as a schema for a potential market, the sensemaker was able to get a comprehensive understanding of each potential market, and was able to easily compare them later on.

For another example, when researching Obama and McCain's energy policies, User MJ13 recognized that they both expressed opinions on a number of energy-related issues, including:

- Federal Gas Tax Holiday
- Taxing Oil Company Windfall Profits

- Domestic Drilling
- Ethanol Subsidies
- Expanding Nuclear Power
- Coal Plants and Coal-to-Liquid Fuel

This list, in addition to two other items: overview and statements from the candidates' campaign websites become the frame structure that User MJ3 used in discussing the policies of the candidates. Schema induction allows structure mapping and transfer between similar concepts. In most cases, policies of the two candidates were discussed and compared under the same framework.

The items in a schema were often added based on seeing the same concept or idea over and over again at several places. For example, while browsing the Obama Campaign Web site, User MJ4 discovered that *"Um,* ***again,*** *investing money into clean energy..."* Clean energy was seen several times and it was recognized as one of Obama's big ideas for resolving the energy problem.

In a similar way, User MJ13 discovered "increase in gas price" as a common theme related to energy: "*'have recent price increases in gasoline caused any financial hardship?'* ***Here we go again****. And, yeah, of course they have."*

User MB5 identified "traditional" as one of Trident's problems: *"see,* ***this is exactly what it was saying before****, about where Trident is really losing some control of the market, because you know, just too traditional for today's youth, I guess.*

As the above examples show, schema induction was often done by seeing the same concept or idea repeatedly. Such concepts or ideas may occur in several sources and possibly expressed differently. For the sensemaker to identify such schema, he needed to be able to recognize that the different expressions were really about the same thing.

Schema induction contributed to structure building by adding the induced schemas. For example, User MJ13 came up with the concepts “energy rebate”, “taxing oil profits of American companies”, and “oil from Strategic Energy Reserve” after reading them in several articles but in different words. For another example, user MB4 added “timing and accuracy” and “rapid turnover” to characterize the rental business. These schemas were induced from particular facts about the concepts.

6.1.2 Mechanisms for Examining Individual Concepts

This section focuses on definition and specification as mechanisms that were used in examining a concept on its own. Other mechanisms (especially those listed under 6.1.3) may also be used for examining concepts, particularly as they related to other concepts. Mechanisms used in examining concepts were often used in a top-down manner and were mostly logic- or structure-driven.

6.1.2.1 Definition

Definition is used to explain or identify different aspects of a concept, such as purpose, function and use. This allows the sensemaker to develop a comprehensive view of a concept.

Definition also helps sensemakers to clarify unclear concepts, to the sensemaker and to the potential audience of any work product that the sensemaking processes produce. When User MJ14 discovered the concept “tax incentive” and decided that it should be included in her story, her think-aloud protocol indicated that she was not clear about what exactly “tax incentive” was, so she went on and created a node in her map labeled "what is it?" and explained it in her story before moving on to talking about the bailout plan. (MJ14)

Definition may also help sensemakers to recognize related concepts. For example, by reading the case description. User MB4 discovered that the company she was researching was "*on the higher end*", she further defined that this "*means they have higher quality dishes, glass, and they can be in like fancy weddings and nice hotels."* Two related concepts of "being on the higher end" have to do with the *quality of their equipment* and possible *occasions* where they can rent this equipment to.

Definition as a cognitive mechanism was not used as extensively as some other mechanisms, partially because the tasks of writing a news article and analyzing a business case often did not involve a lot of new concepts that needed to be defined by the sensemaker.

6.1.2.2 Specification

Specification was used to explicitly state details about concepts and relationships. This was used as the counterpart (or reverse) mechanism of generalization (see Section 6.1.1.5).

There were two main types of specification:

Specify a concept or category with instances or examples. For example, User MB4 specified different types of customers of an equipment rental company by creating nodes linked to the concept "*Customers*", including "*Caterers*", "*Hotels*", and "*Event Planners*". In the same manner she specified "*Competitors*" as including "*Other firms in Boston Area*" and "*Caterers who have their own equipment*".

Specify a claim or principle with examples and particulars. For example, while working on the election news story, user MJ13 specified a general claim that "*Republicans want to 'reverse' their stereotypes*" by an example of "*John McCain saying that small cars are good*." For another example, User MB6 specified several reasons to enter the Hispanic market. *"They're regional. So because it's regional, 'cater more to just this market', 'high growth', and 'Focus on low cost'."*

Several users with news writing tasks created a node labeled "lede" or other similar labels to represent the lede of their news story. They then specified with particulars what the lede was about, for example: "tax incentive", "political debates", and "where does energy stand among the people" for User MJ14. Similarly User MJ9 asked the question "what issues am I going to talk about?" and then specified that they should include "offshore drilling", "nuclear energy", and "reduced gas tax".

6.1.3 Mechanisms Used for Examining Concepts or Relationships

Although examination of relationships cannot be conducted without some consideration of concepts, mechanisms in this category focus on how concepts relate to each other, rather than on the concept themselves. Comparison (of similarity and differences) is a fundamental mechanism for other higher level mechanisms such as analogy (which is based largely on relational similarity) stereotyping (which has to do with the typicality of an object in a category), and classification (which uses resemblance as one of its criteria). Other mechanisms seen to be useful in examining relationships include explanation and inference.

6.1.3.1 Comparison

Items (concepts and facts) that belong to the same category are often compared for their similarities and differences. Comparison was one of the most frequently and successfully used mechanisms.

Comparison was often done between two items, but sometimes users compared more than two items. For example, a sensemaker (MB1) compared compensation options of a retirement plan including stock options, grants, and salary.

The process of comparing two or more items in the same category required careful examination of the character or qualities of the attributes of each item to be compared in order to discover any resemblances or difference. The ability to recognize similarity and to differentiate helps sensemaking in the following ways:

First, comparison for similarity assists recognition of patterns. Several sensemakers with the news writing task compared Obama and McCain's positions on different issues in regards to their energy plan, their articulation of their policy, and even the organizations of their websites. For example, user MJ13 recognized that both candidates had similar plans while she was examining Obama's Web site after she had already investigated McCain's energy plan.

"They've got almost ***the same type of pictures****. Of windmills and trees. <laugh /> "Reduce emissions" from cap and trade!* ***Same thing*** *as McCain. By 2050,* ***also the same thing*** *as McCain."*

Second, comparison helps recognition of new concepts. For example, the same user recognized "*bi-partisanism*" on the environment issue, which became a major theme in her story. *"The environment causes* ***bi-partisanism*** *because they both* *have similar views on the cap-and-trade and on fuel efficiency. And everyone, I mean, no one wants to destroy the environment."*

Third, comparison aids the recognition of important concepts. For example, a user compared the importance of the entities involved in an equipment rental business and noticed that *"'drivers' and 'customers' are the two 'big things' in the equipment rental business".* She then enlarged the nodes representing these two concepts to reflect the result of the comparison.

Fourth, the effort to differentiate among similar things forces sensemakers to conduct more in-depth research and see through the surface. "*So, I think it's just a matter of **looking more in-depth** and seeing specifics on which party, which candidate would really do more for the environment. But on the offset, they both want the same things. They want, you know, wind and solar energy and less reliance on foreign oil, cap-and-trade systems."*

Fifth, comparison between newly acquired materials to prior knowledge allows sensemakers to recognize conflicts. *"See, this is interesting. Some of these articles that we've- that I've found are **different** from what we found before, that said that Trident was **second to Orbit** in the sugarless gum market, because here it's saying that Trident is **number one**."* The user (MB5) did not try to resolve this conflict by settling on one or the other. He created a note page titled "Facts/Data" to store/record these conflicting pieces of evidence.

Sensemakers used various ways to record the result of comparison for later use. In addition to using bigger nodes and bold fonts for important concepts in concept map (MB4), and creating note pages to record these conflicting evidences

(MB5), they sometimes created summary notes to highlight the differences and similarities such as the following note that reflected a comparison of the two candidates' positions on a few energy issues:

- Alternative Fuels: Nuclear, Wind, Biofuels (Obama supported)
- Offshore Drilling (McCain supported)
- Price of Gas (Both candidates)

6.1.3.2 Analogy and metaphor

Analogical reasoning involves the transfer of relational information from a familiar domain (base domain) to the domain to be explained (target domain) (Vosniadou and Ortony 1989). Vosniadou (1989) distinguished two types of analogy: between-domain analogy and within-domain analogy. Most everyday analogies exist between two remote conceptual domains in which the concepts were fundamentally different.

Another advantage of using analogy is that it helps sensemakers to focus on the relational commonalities independently of the concepts in which the relationships were embedded. When examining the youth and adult markets for a gum product, a user (MB5) used similar structures in addressing each: a characteristic of a market followed by a marketing strategy that spoke to that characteristic. The characteristics of the youth and adult markets may have been different (health factor vs. coolness factor), but focusing on the relationship between a strategy and the characteristic it was addressing allowed the sensemaker to construct strategies that were highly efficient and goal-oriented.

In this study, no uses of cross-domain analogies were found, partially because the task domain was not much far removed from everyday life and no such analogies were needed to understand a particular concept or relationship. Likewise, the use of metaphor, that is often used in areas that are abstract such as science education (Vosniadou and Ortony 1989), was not present in the cases of this research.

6.1.3.3 Stereotyping

Sometimes sensemakers made stereotypes of something conforming to a fixed or general pattern. Stereotyping required less mental effort from the sensemakers, although these stereotypes may have reflected an oversimplified opinion or uncritical judgment. For example, user MJ13 made the following stereotype for the Democratic Party on the issue of energy and environment: "*Democrats. Stereotypes: 'cling to'. They want to probably would just be excited that people see them as the more environmentally friendly party."*

Stereotyping helps to explain and connect to prior knowledge and belief. For example, when the same user found that *"Obama wants short-term relief to American families"*, she thought it *"makes sense because he's a Democrat".*

On the one hand, stereotyping largely reduced the mental effort that is needed for thoroughly examining a member of a group by labeling the member as typical of that group, which may be necessary for some sensemaking tasks. On the other hand, it poses the danger of oversimplification.

Stereotyping is related to generalization. The following example is a generalization from a sample to a population: *"Hispanics and blacks are less likely than whites and Asians to own a cell phone. The cause of the lower rate of ownership* *is tied partly to income."* (MB6) Some stereotypes may come from generalizations like this.

6.1.3.4 Classification

Classification was used as one way to organize the conceptual space by relating a concept to a broader conceptual category and grouping of sufficiently alike concepts. For example, users created note pages to keep groups of notes that belonged to the same concept category. The benefit of classification is ease of manipulation and retrieval (Sokal 1974). Users were able to put useful information from articles they found into different categories and retrieve them later on.

Classification was often used at the stage of collecting information and taking notes for the purpose of organizing the notes. Sometimes it was also used when users were creating concept maps to describe the structure and relationship of the objects to each other and to similar objects. For example, in one user's (MJ3) classification scheme shown on her concept map, "issues of energy" included "Domestic Drilling", "Expanding Nuclear Power", and "Coal Plants and Coal-to-Liquid Fuel".

One common feature of classification for the tasks in this study was that the classification schemes were not elaborate, sophisticated, or even complete. They were often just small pieces collected from a bigger classification found in the sources that was not explicitly shown. There might have been multiple classification segments based on different classification criteria. The segments provided flexibility when using in other context. For example, the small classification example above can be easily put into a story map, which is organized by the flow of the news story. In fact, several users combined such classifications with the organization of their task output (a news story or a business plan).

6.1.3.5 Explanation-based Mechanisms

Explanation helped sensemakers to understand causal relationships between concepts and events. For example, User MB1 created causal links between nodes in his concept map for a P&G retirement investment diversification case: *"PG ESOP Debacle" causes its stocks to rise in the 80s and bottom out in 2000. And this further causes demising of value of the employees' retirement plans and PG's effort to diversity employees' retirement options."* By explicitly marking these casual relationships between a sequence of events, the sensemaker was able to understand the cause and effect of the entire matter.

Explanation-based mechanisms also helped the sensemaker use prior knowledge to interpret or explain the newly acquired information. For example, when researching the products of Cadbury (Trident being one of them), User MB5 said in his think-aloud protocol: *"Cadbury could be getting hurt, **because** they just have so many different product lines, um... All these different gums, you know, is it really worth it to be advertising for competing gums?"* The knowledge about benefits and risk of advertising for competing products was part of the sensemaker's prior knowledge, and User MB5 drew a casual relationship between having so many different product lines and potential risk in terms of advertisement effort.

Explanation-based mechanisms were also used in explaining or justifying the claims that a sensemaker made in her write-up. For example, User MJ13 made a claim that *"people are probably blaming the Republicans which would definitely* *throw the outcome of this election,"* and she further explained *"**because** as seen in the other polls it's putting financial hardship on people and people are linking McCain to Bush more and more every day."* In this example the sensemaker used facts about financial hardship and the Bush administration learned from some polls to back up her earlier claim.

6.1.3.6 Inference

Inference is the act of passing from one proposition, statement, or judgment considered as true to another whose truth is believed to follow from that of the former. As *Encyclopædia Britannica* defines it (2009), the derivation of conclusions from given premises can be done by deduction, induction, probability or statistical reasoning. In this study, there was no evidence of probability or statistical inference present because of the nature of the tasks. See section 6.1.1.6 for schema induction.

Two major types of inference were seen: strict logical inference and human inference:

Inferring consequences by strict logic requires a premise that is assumed to be true—a rule—and the conclusion to be derived (Lohman 2005). In many cases, the rule was implicit. For example, after reading some facts about more young people are chewing gum than American adults by certain percentage, User MB5 said in his think-aloud *"So as a result of this Trident must create fun, witty advertisements focusing on the youth market…"* The implicit rule behind this inference is that younger people like fun, witty advertisements. Representing the inference by deductive reasoning:

Premise/Grounds: Younger audience is a major consumer of gum.

Rule/Warrant: Younger people like fun, witty ads.

\---

Conclusion/Claim: Therefore Trident must "create fun, witty advertisements for the youth market".

Box 6-1: Logical Inference Example

Another type of inference draws judgments or conclusions based on circumstantial evidence. This type of inference, which has been mostly studied by cognitive scientists, is not as strict and often introduces inferential errors (Nisbett and Ross 1980). This type of inference was more common in the sensemaking cases. Though the conclusions were often not logically derivable from the assumed premises or facts, they still possessed some degree of probability relative to the facts.

For example, User MB5 inferred from the fact that *"it doesn't seem like people really prefer the sugar"* to the conclusion that *"this could be part of the differentiation just to advertise or to really point out the sugarless fact, to try to sell that"*. Different from logical inferences, the conclusion cannot be guaranteed to be true even if the premise is true:

1. There was no such rule that guarantees selling sugarless gum would lead to a successful differentiation strategy when people do not seem to prefer sugar.
2. The evidence was circumstantial: people not preferring sugar does not mean that people prefer sugarless gum.

However, the conclusion was still very likely to be a valid one. It was reasonable to think that since people do not want too much sugar intake, they may prefer sugarless gum, and it would be a reasonable strategy to advertise the sugarless aspect.

6.1.4 Mechanisms Used for Examining Anomalies and Inconsistencies

Inconsistency may exist at three levels (data, schema, between data and schema) in prior knowledge, new information, and between prior knowledge and new information. The following table summarizes these inconsistent situations:

Table 6-3: Detection of Inconsistencies

	Prior knowledge	**New information**	**Between prior knowledge and new information**
Data level	Inconsistent facts in prior knowledge	Inconsistent facts in new information	Facts found in new information is inconsistent with facts already known/collected
Schema level	Conflicting schemas in prior knowledge	Conflicting schemas in new information	Schema detected in new information conflicts with schema established in prior knowledge
Between data and schema	Data and schema in prior knowledge do not fit	Data and schema in new information do not fit	Facts known do not fit newly discovered schema, or schema established cannot account for new facts

The inconsistencies, once solved, often moved the sensemaking to the next level. Mechanisms that were used in examining anomalies and inconsistencies included elimination, semantic fit, and Socratic dialogue.

6.1.4.1 Elimination

Elimination was commonly used to exclude facts, concepts, or relationships that were not applicable. For example, User MB5 deleted notes talking about a very specific target user group, "smokers", because it is *" just too specific to market on"* which is not consistent with Trident's overall marketing strategy. Very similarly User MB6 eliminated "Alaska native" from the target audience of a wireless phone company because *"that is such a small population, we're not even- we're just going to completely delete this. There is no way that they would be our target market"*.

In other cases, sensemakers eliminated facts that show conflicting evidence *"… I wanted that article to say something else. I have to disregard it."* This is one of the four ways to resolve inconsistencies and conflicts (See Section 5.2.3).

6.1.4.2 Semantic fit

Concepts do not exist in isolation and will only make sense when they are connected to other concepts to construct the meaning of a whole. Sensemakers used semantic fit to examine the reasonableness of a fact, concept, or relationship as it related to the meaning of other concepts in the knowledge structure.

For example, User MJ9 found an article talking about Americans "showing stronger support for pro-environment policies when economic conditions are considered good". She further examined the fitness of this fact to the structure of her election story as a whole: *"See, in cases of crisis, I don't know if this is relevant."* Since the election was taking place during an economic melt-down, the sensemaker was not sure whether this fact really fit her story. User MJ9 did not recognize that from the article one could infer that when economic conditions are poor support for pro-environment policies will be weak.

For another example, when examining a poll that said “a lot of people are saying it (gas price) affects them”, User MJ13 thought this fact might have to do with the election, though not directly. In other words, it fit the story only tangentially. She might still have copied the polling results in her notes, but it was not very likely to end up in her story.

The reasonableness of fit may be related to the knowledge structure as a whole, or it may be related to other sporadic elements of the knowledge structure. In another example, User MJ1 was compiling notes from several interviews she did previously on an election news story (a different task from the election and energy news). When examining one interview, she said: *“yeah, her main issue was abortion… I don’t know if I should include her* ***in this story,*** *or if it’s worth typing up all of her information, because her interview wasn’t as rational as* ***the other ones,*** *I guess. But maybe her point of view is worth sharing. We’ll see.”* Comparing the interview to the other interviews and considering the story as a whole, the sensemaker examined the reasonableness of this interview data in relation to other elements involved in the story, although she did not come to a conclusion about whether to include this particular interview or not.

In some cases, sensemakers examined the fitness of a fact, concept, or relationship, realized that it might not fit perfectly, but were still satisfied to keep it where it was put temporarily. For example, User MB5 created a node for “needs”.

He was not sure "if this really fits on the map", but decided to leave it for the time being.

6.1.4.3 Socratic dialogue

Sensemakers conducted critical dialogues to facilitate the awareness of inconsistencies in the current schema. Some sensemakers conducted dialogues with themselves, reflecting on the knowledge structures and asking critical questions about different parts involved. Socratic dialogue allowed for in-depth understanding of various issues concerning the sensemaking task.

For example, when browsing a poll from the Pew Research Web site – policy priorities for the country regarding development of new energy sources – User MJ13 asked the questions *"are they really caring about the environment? Or do they not want to run out of oil and pay a lot for it? <laugh />"* Such questions allowed the sensemaker to gain more in-depth understanding of the issue of developing new energy sources.

In another example, User MB5 asked a question *"So they spent 5 million? Oh, no. Yeah, in 2003, spent about 5 to 10 million dollars on advertising?"* Despite the unusually high amount of money spent in advertisement, User MB5 discovered that it did not bring the company much gain in market share, so the sensemaker concluded *"Wow, this commercial must have set them back."*

Some sensemakers conducted Socratic dialogue when they were confused. For example, when User MJ3 found an article offering a general introduction of the energy issue, she was not sure if this article was useful. "*I definitely will need this page because? I just did not need this page? I just did not need this page? Oh, okay,* *maybe it's good*." When reading about a marketing research methodology User MB3 asked "*what does operational definition mean? 'Statement of precisely which observable characteristics will be measured and the process of assigning a value to the concept.' … still not very clear about operational definition.*"

Socratic dialogue as a cognitive mechanism was not always used in planned and sophisticated ways. It was often conducted in an ad hoc fashion. By asking and answering critical questions, sensemakers were able to look at facts, concepts or relationships more critically, and keep track of their thoughts.

6.2 Data-driven vs. Logic/Structure-driven Mechanisms

As Figure 6-1 suggested, some mechanisms are primarily data-driven, while some are primarily logic/structure-driven. Some mechanisms may be used as both data-driven and logic/structure-driven, and some do not belong to either category. The following table summarizes the mechanisms in terms of their inclination toward data-driven or logic-driven:

Table 6-4: Cognitive Mechanisms by Data-driven vs. Logic-driven

Data-driven	**Logic-driven**	**Both**	**Neither**
• Key item extraction • Restatement • Summarization • Judgment / evaluation • Generalization • Schema induction	• Definition • Specification • Elimination • Explanation-based mechanisms • Inference	• Comparison • Semantic fit • Analogy • Stereotyping • Classification	• Socratic dialog

The distinction between data-driven and logic-driven is not always clear. Some mechanisms may be based more on data than on logic and vice versa. A better way to view these mechanisms from the data- vs. logic-driven perspective is to view them as being on a continuum with data-driven and logic-driven ends. The relative position of a cognitive mechanism on the X-axis (Figure 6-2) indicates their inclination toward the data-driven or logic driven end.

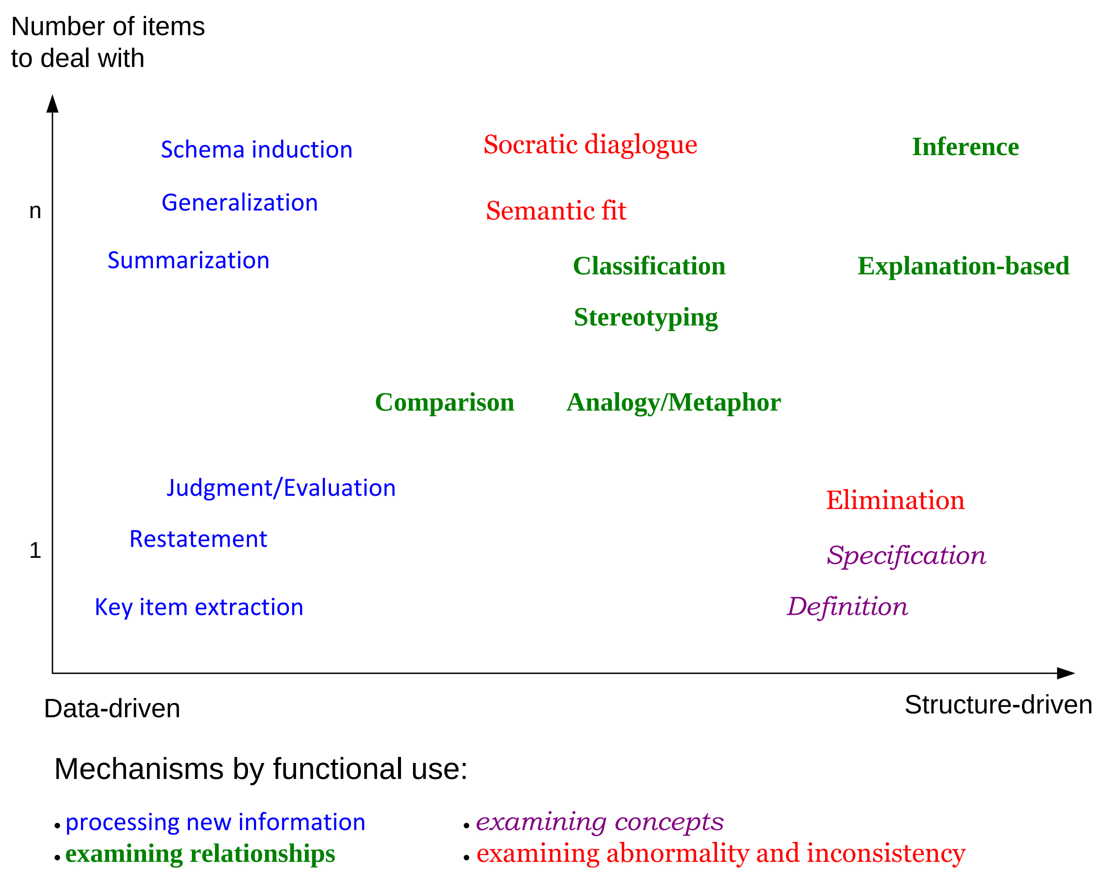


Figure 6-2: Cognitive Mechanisms, Repeated here for ease of reference

6.2.1 Data-driven Mechanisms

The mechanisms that belong to the functional category for processing new information are mostly data-driven. The mechanisms were applied to raw data / facts, and schemas were built.

Key item extraction can be used purely data-driven and it can be used to instantiated structure. It can only be done after seeing the item to be extracted from data. Therefore it was always driven by data. This is not to say that key item extraction does not require any logical thinking. For example, while a user (MJ14) read the article "Energy tax credits gain momentum from bailout," from Reuters.com, she said *"'700 billion dollar economic rescue package', so now it's attached… the (energy) tax credit at one point..."* The sensemaker extracted the item "Seven hundred billion dollar economic rescue package" because it connected to something she saw earlier: the energy tax credit. There was logical and analytical thinking going on, but the data triggered this cognitive mechanism.

Restatement and *summarization* were also used on the data found. Sensemakers restated or summarized what they read to make it more accessible to them. Similar to key item extraction, this can only be done after seeing the data. For example, A sensemaker (MB5) restated a long sentence about contributing factors of a market growth as *"that's what I was saying how sugarless gum is the new product in the category."* No restatements or summaries existed without the original, and they were always driven by data.

Judgment/evaluation, generalization, and *schema induction* may require some low degree of knowledge about structure or logic but they are still largely driven by new data; the primary goal of such mechanisms is to derive some structure from the data.

For example, a sensemaker (MJ13) judged the *“rationing gasoline and oil”* as *“obviously ridiculous”*; to do this the sensemaker had to have some knowledge about what “rationing gasoline and oil” is, how it relates to the energy issue in general, and why it is ridiculous. But this mechanism was driven by seeing a poll about “rationing gasoline and oil”.

For another example, the same sensemaker generalized from a poll that “*the entire country – bipartisanly – wants to fix the environment”*. This required the sensemaker to have some knowledge about how representative the poll was. But the generalization was still data-driven because it was based on a specific fact that the sensemaker learned.

Other mechanisms under this category often followed “key item extraction”. For example, sensemakers used a “key item extraction – summarization” combination to write a summary regarding important concepts. For another example, a “key item-extraction – judgment” combination was used to express opinions on the main concepts mentioned in an article. Table 6-5 summarizes the occurrence of these data-driven mechanisms following “key item extraction”.

Table 6-5: Occurrences of Data-driven Mechanisms

	All	Preceded by "key item extraction" in close proximity		Preceded other mechanisms	
		N	%	N	%
Key item extraction	337	289	88%	48	12%
Restatement	67	41	61%	26	39%
Judgment or Evaluation	55	37	67%	18	33%
Summarization	64	26	41%	38	59%
Generalization	16	10	63%	6	37%
Schema induction	23	10	43%	13	57%

Key item extraction was the most-used data-driven mechanisms. In average, 55% of the occurrences of other data-driven mechanisms were used following the use of "key item extraction".

When combining key item extraction with restatement, judgment or evaluation, and summarization, users often used key item extraction as a filtering mechanism to get the target concept for further processing using the other mechanisms. Examples of the combinations were illustrated below:

- *Key item extraction – restatement:* User MB5 quickly browsed a section and extracted the named entity "Orbit" as a competitor of the gum company that he was researching. Then he read the sentence *"After a couple of decades of success in the UK, sugarless gum brand Orbit launched in the US in 2001 and quickly established a place for itself among the category leaders"* and restated this fact as *"Orbit just* *has only been in the US since 2001. Wow, that's seen huge growth. Um, (it) was in the UK, that's amazing."*

- *Key item extraction – judgment/evaluation*: while reading the polling data, User MJ13 first extracted an important issue from the poll *"increasing federal tax*", read the results "*opposed!*", and then expressed her opinion on this issue *"No one wants to pay more money so obviously everyone just wants to blame somebody. They don't want to have to pay for it."* The key item extraction was again used as a filtering mechanism where only the important facts, concepts, or relationships got judged.
- *Key item extraction – summarization:* after reading several paragraphs about McCain's energy plan, user MJ13 summarized *"Ok, so clearly… he (McCain) wants to break dependency on foreign oil like everyone knows."* This summary was based on extracting a key concept *"oil dependency"*.

The combination of key item extraction with generalization and schema induction suggested a different pattern of use. Often several individual facts had to be extracted before a generalized claim can be made or a schema can be induced. For example, User MB5 saw several similar facts about the use of gum in adult and youth audiences, then generalized some trends or themes (adults prefer sugarless and youths prefer bold flavor). These cases of key item extraction may not happen successively in a row. Often key items were extracted from several articles over a period of research with other activities in between. The fact that the same items were extracted over and over again from different sources made generalization and schema induction possible.

While in most cases key item extraction was data-driven, using key item extraction to instantiate structure was actually logic-driven. It was not possible to determine for every case that used key item extraction mechanism whether it was driven by data or logic, because users simply did not verbalize every trivial thought they often spoke only those seemed important to them. A future research direction could be a refined analysis of the use of key item extraction mechanism.

6.2.2 Logic- or Structure-Driven Mechanisms

Logic-driven mechanisms were used in examining structural elements including concepts and relationships.

Definition was used to explain or identify different aspects of a concept, such as purpose, function, and use. Such aspects represented the structural dimensions of a concept, which were independent from any specific data on those aspects. Of course facts needed to be added to the slots to give complete definition, but defining a concept was driven by logic or structure.

Specification is structure-driven in a similar way to definition. Specification has to include facts to be complete, but it was primarily concerned with concepts and general claims rather than specific data. To specify concepts with examples or to specify a general claim with facts, the primary goal is to expand the structure with details.

Elimination functioned based on elimination criteria – a set of rules that excluded certain structure elements or facts from the knowledge schema of the task.

For example, a sensemaker decided that she was not going to talk about "global warming" in her story (MJ3) because her knowledge was best suitable for an overview of the energy issue instead of focusing on a particular aspect. She eliminated this structure element (the "global warming" concept) based on the consideration of the story structure as a whole.

Explanation-based mechanisms were used to make clear the cause or reason of some observations. It required the sensemaker to use logic to interpret or explain the causal relationships between two events. For example, when User MB5 claimed that *"Cadbury could be getting hurt,* ***because*** *they just have so many different product lines, um... All these different gums, you know, is it really worth it to be advertising for competing gums?"* The user's underlying rationale that drove this claim appears to be that advertising for competing product lines could possibly hurt a company's advertisement effort.

Inference was on the logic-driven end of the continuum. In fact it is the process of deriving logical consequences of assumed premises. It cannot function without logic. The rules for inference may be implicit as in the example shown in Box 6-2:

Premise/Grounds: Younger audience is a major consumer of gum.

Rule/Warrant: Younger people like fun, witty ads.

Conclusion/Claim: Therefore Trident must "create fun, witty advertisements for the youth market".

Box 6-2: Logical Inference Example, Repeated for Ease of Reference

In the above example, the sensemaker (User MB5) did not explicitly state the rule for deriving the conclusion "therefore Trident must 'create fun, witty advertisements for the youth market' " based on the fact that youth audience consume more gum than the adult audience, but it was the underlying driving force for getting to the conclusion.

6.2.3 Cognitive Mechanisms Spanning Both Categories

Mechanisms in this category can be both data-driven and structure driven. These mechanisms may be performed on data or on structure elements.

Comparison was used by sensemakers to identify the similarities and differences of items that belong to the same category. For example, a sensemaker compared the market share of two companies (data), and he also compared two potential markets for different aspects such as age group, focus, and product line (structure). User MB6 built a schema for ethic groups as potential markets for a mobile phone company, and compared the groups using that schema. Comparison may be driven by seeing a particular fact and recalling a similar or different fact, or it may be driven by thinking and analyzing two concepts logically. Sometimes comparison at the fact level may lead to comparison at the structure level and vice versa. For example, by seeing similar facts on different energy issues such as "…windmills and trees" "…reduce emissions from cap and trade" and "…by 2050", a sensemaker (MJ13) concluded that *"they've got almost the same type of pictures"*, meaning that they have similar positions on various energy-related issues. The comparison of the overall energy policy is a structure-level comparison.

Analogy was more logic / structure driven than comparison. The relational information to be transferred often has to do with structures. Analogy may be recognized by seeing similar properties of concepts (facts), but very often it focuses on the relational commonalities independently of the concepts between which the relationships are hold (structure).

Semantic fit may be used in examining whether a particular fact fits a concept or category (data-driven) or in examining whether a concept fit the knowledge structure as a whole (structure-driven). Here is a data-driven example: User MJ9 examined the fitness of a fact "Americans show stronger support for pro-environment policies when economic conditions are considered good" but since the election was taking place during an economic melt-down, the sensemaker was not sure whether this fact really fitted her story. For a structure-driven example, User MB5 examined how the concept "needs" fit in his concept map as a whole and concluded it did not really fit the map.

Stereotyping was sometimes created by a sensemaker to demonstrate that a particular fact conformed to a fixed or general pattern (data-driven). Sometimes stereotypes were used to explain how things were related (logic-driven). For example, User MJ13 put down "cling to" energy as the Democrats' stereotype on this issue and *"they want to probably would just be excited that people see them as the more environmentally friendly party."*

Classification was used as one way to organize the conceptual space by relating a concept to a broader conceptual category and by grouping of sufficiently alike concepts (logic or structure driven). Sometimes users also classified facts into the conceptual categories (data-driven). See Section 6.1.3.4 for examples.

6.2.4 Neither

Socratic dialogue was not easily identified as data- or structure-driven. Occasionally sensemakers asked questions about a particular fact (*"So they spent 5 million? Oh, no. Yeah, in 2003, spent about 5 to 10 million dollars on advertising?"*) or about the knowledge structure (*"do I need this concept?"*), but very often sensemakers asked questions when they lost track of where they were in the sensemaking process or when they were confused. It was not clear whether such thought process was driven by data or logic.

6.3 Overall Sensemaking Approach and Cognitive Mechanisms

Sensemakers have different ways to approach sensemaking tasks. A sensemaker might use a predominantly bottom-up approach where she started with little previous structure and derived all structure from the data, or a sensemaker could use a predominantly top-down approach where she had a pre-defined structure to build the structure at the beginning of the task by analyzing the task and then searching for data to instantiate the structure. More often, sensemakers used a mixed or combined approach where they started with some pre-defined structure, or built part of the structure by task analysis and then derived other parts of the structure from data as they searched for data to instantiate the structure.

It seems reasonable to assume that sensemakers with a top-down approach may use more logic-driven mechanisms, and sensemakers with a bottom-up approach may use more bottom-up mechanisms. To see whether the overall approach of the sensemaker (top-down or bottom-up) influenced their use of cognitive mechanisms (data-driven vs. structure or logic-driven), the researcher assigned a score for the overall approach from a five-point scale for each case:

Top-down		Mixed		Bottom-up
1	2	3	4	5

The overall score of a sensemaking approach was determined considering the following evidence:

- How much structure was built up-front before doing any search
- Sources of the concepts and relationships: whether they were from new material or prior knowledge and analysis
- The sensemaker's description of his/her sensemaking approach in the post-session interview

Table 6-6 shows the overall approach of the sensemakers and their use of data-driven vs. structure- or logic-driven mechanisms, including the number of instances of data-driven mechanisms and number of instances of logic driven mechanisms.

Table 6-6: Sensemaking Approach and Use of Cognitive Mechanisms

Over all approach	**Case No.**	**No. of instances of data-driven mechanisms**	**No. of instances of logic-driven mechanisms**	**Ratio =** $\frac{\text{No.of data-driven instances}}{\text{No.of logic-driven instances}}$
Top-down (1)	MJ4	10	3	3.3
	MJ5	12	2	6.0
	MJ8	22	4	5.5
Mixed (2)	MB5C2	73	15	4.9
	MB6	26	5	5.2
	MJ13	144	22	6.5
Mixed (3)	MB5C1	90	34	2.6
	MJ3	54	11	4.9
	MJ9	29	13	2.2
Mixed (4)	MJ1	31	3	10.3
Bottom-up (5)	MB1C1	42	7	6.0
	MB1C2	10	0	-
	MB3	25	2	12.5
	MB4	50	11	4.5
	MJ14	22	4	5.5
	MJ15	7	4	1.8
Average		40.4	8.8	4.6

Correlation analysis suggested that the overall approach of a sensemaker did not seem to influence the ratio of instances of data-driven and logic-driven mechanisms used in sensemaking.

The analysis suggested that regardless of whether the overall sensemaking approach is top-down or bottom-up, a sensemaker may use both data-driven and structure-driven mechanisms for his sensemaking task. The following table summarizes the situations in which structure-driven and data-driven mechanisms were used in bottom-up or top-down approaches.

Table 6-7: Overall Sensemaking Approach and Step-level Cognitive Mechanism

Overall approach / Mechanisms used in one step	Bottom-up	Top-down
Data-driven	1	3
Structure-driven	2	4

1. If a sensemaker's overall approach was bottom-up, he used data-driven mechanisms such as key item extraction, comparison, generalization, and so on to come up with higher level structural representations. This is intuitive and easy to understand.
2. When the overall approach was bottom-up, a sensemaker may still have used logic-driven mechanisms. For example, User MB1 used explanation-based mechanism (logic-driven) to examine the casual relationships among sequences of events (data) he discovered from the case material.
3. When the overall approach was top-down, most users still used data-driven mechanisms such as key item extraction, restatement and

summarization to process new material when trying to instantiate the established structure.

4. If a sensemaker's overall approach was top-down, he started with higher level structures and used logic-driven mechanisms such as definition, specification, explanation to expand the structure and to get to detailed data.

To summarize, users used a variety of data-driven and structure-driven mechanisms to process new information, to examine concepts and relationships, and to detect anomalies and inconsistencies. The use of data-driven and structure-driven mechanisms was not influenced by whether the sensemaker used a bottom-up or top-down approach.

Chapter 7: Implications for Design

Design implications for tools that assist sensemaking tasks emerged through the analysis of the think-aloud protocols and screen recording data using the proposed sensemaking model and through interviews with participants. The details in the proposed model provide a better basis for designing sensemaking support systems. The model captures the task level iterations and the step-level cognitive mechanisms. Although there were various tools that may assist some aspect of sensemaking (reviewed in Section 2.5), very few tools were designed to assist sensemaking in particular. This chapter discusses the implications for the design of a sensemaking tool (or tools) that focuses on the following aspects:

1. Representation and manipulation of the conceptual space
2. Assistance at the cognitive mechanism level
3. Assistance at the task level, especially in building and instantiating structure
4. A design framework of functionalities and data architecture

This thesis does not intend to claim that an information system that aims to assist sensemaking should incorporate all possible features and techniques discussed in this chapter. Rather this chapter aims to provide a design framework for information systems that assist sensemaking tasks. The design of any system should consider the nature of the sensemaking task(s) it aims to facilitate and the characteristics of the intended user group.

7.1 Representation and Manipulation of the Conceptual Space

When a sensemaking task is difficult, sensemakers use external representations to store information for repeated manipulation and visualization (Stefik, Baldonado et al. 1999). The conceptual space is composed of sensemaking artifacts in different forms of representation. Research in visualization (Gaines and Shaw 1995; Chi and Card 1999; Card 2009) has put much emphasis on visualizing the collection that users search rather than the conceptual structures that users create through sensemaking tasks. The model proposed in this dissertation provides the basis for designing and evaluating tools that help structure the representations in a sense-maker's conceptual space to provide better sensemaking support to information system users.

7.1.1 Multiple Representations of an Underlying Structure

The idea that the system should provide multiple representations to accommodate different users and users at different task stage is well-established (Minsky 1975; Ingwersen 1992; Ingwersen and Järvelin 2005). The present study sheds light on how multiple representations may help sensemakers in accomplishing sensemaking tasks in particular.

Multiple representations of the same underlying structure may include network representations such as maps (spatial), concept hierarchies such as an outline or directory, frames, and text representations (Flower and Hayes 1981; Fikes 1985; Sowa 2006). Each form of representation offers different contributions to users' sensemaking. Users mentioned that detailed views of notes were most useful at the product creation stage when writing the news articles and when answering detailed case questions. Detailed views of notes in a particular section or note page were also useful when a sensemaker worked on a particular concept.

The note structure / outline helped users to have easy access to their notes, particularly if they were *"reading something that had a lot of information to remember"*, being able to have some structure (directory or outline) users were "*able to sort it into what is important".* (MB6)

The note structure also helped at times when the sensemaker wanted to get an overall picture of the notes and the task, or when a user "*forgets what I have where I cannot see it all at the same time.*" (MB1)

Concept maps were more useful *"for understanding the relationship between everyone"*, for example, in a case *"there was the company and then the people in the company, the customers, competitors, the bank, and they are all different part of the case so that (map view) was useful."* (MB4)

Multiple representations may be especially helpful in collaborative sensemaking where a single representation form may not be best suited to the styles of all collaborators. If collaboration is asynchronous, sensemaking handoffs are critical to the success of sensemaking (Sharma 2007). Handoffs that can be viewed and manipulated in multiple representations may be able to compensate for the differences between collaborators to some extent.

Every representation should be linked to an underlying structure, and the changes in representation may be reflected in the underlying structure or structures. The options of how the underlying structure(s) may be changed in correspondence to the changes in representations include:

- One underlying structure – if one representation changes the structure, the other representation(s) is changed to reflect the change.
- Multiple structures – change in one representation does not necessarily result in changes in other representations. The system could keep an underlying structure for each representation or keep a shared underlying structure for selected representations specified by the sensemaker.

There should be multiple levels of control. Users should be able to specify how they would like the representations to be updated, whether the unit of specification is at the level of each individual change or at the level of a representation format. For example, a user may choose to automatically update concept maps based on the note structure, whereas a user may choose not to update concept maps when notes are updated. A user may choose to be notified of every change in structure and decide for himself whether he wants a particular change to be reflected in another representation.

With the underlying structure(s) for multiple representation formats, the system could provide users with tools for specific tasks. For example, the system may provide an outlining tool that transforms the concept maps to outlines with nodes (concepts) as headings of the outline, and with relationships and notes listed under the related concepts.

7.2.2 Text Representations

Expressed with written natural language, text representations can be more precise, nuanced and detailed than graphical representations. The production of a text representation (the process of writing) is an important step in producing and combining structure and data. Text representations may include forms of pure texts and texts with formatting.

Suthers (1998; Suthers 2001) argues that text representations often fail to represent structure and relationships saliently. However, text representations are very expressive and can contain details that no other forms of representation can; moreover, the process of creating a text representation enables reflective and analytical thoughts (Flower and Hayes 1981). As a result, notes were often taken in the form of free text, and very often they were copied and pasted from the original source. This is particularly true for the news writing tasks, because the users had to refer to the original for quotes. Users also wrote free-text notes or annotations to record their analytical thoughts, questions and insights.

Notes served as a data and idea warehouse for the sensemaking task. Most users extracted segments of text with a hyperlink to its original source, while a few also included full-length articles that were particularly useful to them. This note space differed from bookmarks of websites or saved documents since the process of building such notes involved some level of analysis and synthesis. Most users found the note-taking application useful for their sensemaking task. Notes helped sensemakers to streamline their thinking (User MJ1).

With the particular note-taking tool used in this study (MS OneNote), notes are represented as freely movable textboxes, which give the user flexibility in the degree of structure she wants. Some users found that it was *"easier to organize different concepts"* with the ability to "*move them wherever I want and not only in a* *horizontal or vertical way."* (User MB3) *"OneNote was really helpful just being able to drag stuff in there without having to be restricted to the format of a Word document, just how it makes those textboxes."* (User MB5) Some were annoyed by the unaligned textboxes and preferred the more traditional approach where a document starts at the top of a blank page and everything is aligned.

From the analysis of different types of notes as accretion to the conceptual space, it would seem useful for the system to differentiate copied text from user generated text (such as annotations or restatements). Due to the restrictions to link the applications used in the research, users often used placement of an annotation to make such link explicit. This suggests that the system should provide annotations at individual fact level or at structure levels, allowing annotations about particular facts and general observations about structures and themes, as well as creative thoughts of the sensemakers that are not related to any particular data/fact.

Users expressed the need to bring in some structure in text representations so that they could view texts at different levels of abstraction. OneNote provides a three-level organization of notes for each "notebook":

- Notebooks represented as folders on the left panel
- Sections represented as tabs at the top
- Note pages that are listed on the right panel and can be viewed in the main panel

Most users liked the ability to have different sections and pages so that they were able to structure their notes. However, only one note page can be viewed in the main panel at one time. There was a need to provide structure in a single note page by providing "*different titles for different textboxes*" for example (MB4).

On the one hand, text representations in a sensemaking system should allow flexibility in terms of how notes are taken. For example, the system should allow user-defined note formats (freely movable textbox or traditional document structure) and connect notes and annotations at different levels of the conceptual structure. On the other hand, text representations should also bear some structure so that they can be viewed with desired amount of detail.

It seems to be useful if the system can provide some ability to structure text representations:

- Hierarchical structures may be represented as note pages, subpages, and sections; within each level of the hierarchy, the ability to format text representations by levels of headings, fonts, and colors may be useful.
- Graphical layout or arrangement provides method to represent structure in addition to the hierarchy.

Providing pre-defined structure, such as a frame, for notes may be useful to users who work on tasks that are relatively well-defined.

7.2.3 Graphic/Network Representations

Aside from usability issues of a particular software application, graphic/network representations still pose challenges to typical users. Some of the challenges include:

1. Limited forms of representation. With any particular visualization software, the forms of representation are limited.

Users with different types of tasks often need different forms of graphic representation. For example, a concept map may be more suitable for a case involving interrelated players and different aspects to consider but not for a case involving prediction of market share based on computation of financial figures. A challenge for task-independent tools is that they need to support different forms of representation (Wang and Haake 1997). The concept mapping software (CMap) used in this study was a generic tool which does not provide multiple representation forms, and users sometimes wanted to have more forms to explore, such as charts (bubble charts, flow charts) and tables(MB6).

To overcome this challenge, the system should differentiate representations of concepts and relationships that are common to several tasks and representations of concepts and relationships that are specific to a particular domain.

2. Structured information is often only a partial abstraction of the information it represents.

This challenge is not unique to network representation. Structured data loses certain aspects of the information it intends to represent (Hsieh and Shipman 2002). When asked to explain their concept maps to the researcher at the post-session interviews, several sensemakers had to think hard or refer back to their notes occasionally to interpret the maps they had just created. It is reasonable to assume that this would pose even more of a challenge in collaborative environments. Moreover, it may be very difficult to construct partial abstraction. In the case of this research, relationships were especially difficult for users to describe with graphical representations. Users left many relationships unlabeled (CMap put a "????" automatically for unlabeled relationships): sometimes a relationship was too complex to express by a condensed label, and sometimes a relationship was so obvious that it did not need a label.

In some task domains, there may be a common set of relationship types that are likely to apply to many tasks. In this case, giving the types of relationships, perhaps color-coded, may help users with building structure using these relationships. Automatic information extraction to fill such pre-defined relationship types may help users fit data into structure. However, many relationships were specific to a particular case and a particular user construct. For example, in an equipment rental case, the sensemaker constructed the relationships based on "*what they do for each other"* and *"what their function is", "like for this one they give them equipment for events, and then here they give them money to have business, and this is their competition"* (MB4). Some of these relationships may be generalizable, but most are unlikely to apply to another case or another sensemaker. So instead of trying to automatically construct all different types of possible relationships, it may be a good idea to use relationships that are common and easily extracted, and provide better support for manually creating other relationships that are specific to individual cases.

It is important to allow unlabelled links where a relationship exists but cannot be easily defined. Going even further, Relationships may be embedded in the layout of a graph (for example, peer relationship) without explicitly representing links as lines.

3. Extra work for creating network presentation. *"I didn't feel the need or the necessity to create a map."*

Creating the network presentation often requires extra work from the users. Sometimes it may not be worth the extra effort for a small-scale sensemaking task where everything can be handled in the sensemaker's head: *"with a story so short, like 400 words, I can pretty much organize it in my head, and it's just easier to me"* (MJ4). However, users also suggested that such representations would be helpful for more complex tasks: *"had the case been more complex, a visualization tool would be more helpful, because it's easier to visualize everything on CMap"* (MB1). The user suggested that he would lay out the different aspects of the case, including the financials and the competitor information. With the network representation, he can locate information about some aspects quickly by simply clicking on a node. At some point on the continuum of complexity of tasks, the benefit of visually laying out related concepts in a complex task where they cannot all be processed inside one's head could outweigh the cost of creating such representations.

CMap posed some usability issues that contributed to the extra effort. Some are particular to this application, but a more general usability issue for graphical representations includes layout: automatic layout does not always get the concepts in the desired position users designated, and manual layout required a lot of time moving around different parts of a concept map, which can be bothersome. It seems that users could benefit from supports for automatic or computer assisted creation of concept map that are flexible.

7.2 An Architecture that Links Structure, Notes/Data, and Sources

Several users mentioned that they liked the ability of the note-taking tool that kept the source URL or file location automatically when they copied a segment from a particular Web page or file. And they complained about having to take extra steps to link a concept node in CMap to the related notes in OneNote. Some still found it worth the effort to link the two since it was very helpful "*to lay out your ideas and create a web for analysis that includes all sorts of materials from different parts of your computer and pull them together"* (MB1).

What was referred to by this user as "a web for analysis" suggested a simple notion of an architecture that links information objects of all kinds. Such an architecture may be implemented in one integrated system, or it may be implemented in several components with easy and seamless shifts between them.

The sensemaking workspace contains a rich set of information objects which needs to be represented and manipulated by users. The categories of information objects include:

- Structure elements that users created, including concepts, entities, and relationships represented in multiple forms such as nodes and links in graphical representations, headings and subheadings in outlines or folders/pages in directories, or tags assigned to data
- Data, including copied notes and free-text annotations added by users and the links to source
- Sources: original documents that users collected and their source information with time stamps

- Intermediate and final products of sensemaking with links to sources, such as the intermediate knowledge structure composed by structural and data elements, a written report, or a presentation.

None of the above categories exist in isolation. This study suggested that the ability to get to the related elements in one category from another is much needed. When creating their sensemaking tasks, several users referred to their notes either to seek for insights from their annotations or to look for quotations or facts to include as supportive evidence for their claims. For example, while writing the marketing proposal for Trident, User MB5 constantly went back to notes to find specific facts and to look for analytical notes on advertising ideas. Users had to do a lot of copying-and-pasting and going back and forth in order to keep track of an insight or thought. The ability to link the sensemaking product with the data and analytical notes of the sensemaker was very much needed, especially during the product creation stage where thoughts come together and knowledge was updated and reinterpreted with goals of the task.

7.3 Assistance at the Cognitive Mechanism Level

The analysis of the cognitive mechanisms used in sensemaking provides a useful framework for designing information systems. This section describes system functionalities that may assist with tasks performed using these cognitive mechanisms.

7.3.1 Reducing Complexity and Making New Information Accessible

Often sensemakers need to process large quantities of information from multiple sources in a relatively short time. Two major functions of the mechanisms used in processing new information are to reduce complexity and to make the new information more accessible. Information systems in general, not limited to those intended to assist sensemaking, should provide assistance to help users with these goals.

7.3.1.1 Information Extraction

Key item extraction was the mechanism most used by almost all users. It was one way that users dealt with massive information by extracting the key items and limiting the scope of further processing to the items of interest. Many of the items extracted were filtered and only a few become keywords for the next search, headings of a note page, or a concept node in the concept map. Filtering these extracted items is a process that requires human intervention, whereas the extraction of such items from massive text and bringing them to the attention of sensemakers can be done by systems.

Information extraction may help sensemakers in two major ways: automatic information extraction and automatic assistance for manual information extraction. Automatic information extraction techniques have long been studied and named entity and relationship extraction in the domain of people and organizations have yielded somewhat satisfactory results (Cardie 1997; Grishman 1997; Doddington, Mitchell et al. 2004). Information extraction techniques can support sensemakers with “instantiated structure elements” i.e., entities and relationships along with the sources from which they were extracted. For example, the system might suggest preliminary formal statements for users to examine and filter, saving users the time of reading retrieved documents and extracting relationships manually. The system may provide several options for dealing with such extracted entities and relationships once selected by users. For example, a user might right-click on an entity identified by the system, creating a node with the entity in his knowledge structure in any formats described in Section 7.1. If at least one entity of a relationship already exists in the user's knowledge schema, the sensemaker should be able to merge the relationship into the existing structure.

Automatic assistance for manual information extraction may also help users in sensemaking. For example, the system may highlight potentially important information, such as named entities and/or relationships and allow users to seamlessly extract the information to put into notes (such as drag-and-drop).

A key design issue to consider is how to organize and integrate extracted results into the emerging conceptual structure of the sensemaker. The system should present the extracted results in ways that come naturally to the sensemakers, so that information extraction helps reduce the amount of information that users have to process without posing additional cognitive load to sensemakers.

7.3.1.2 Paraphrasing and Summarization

A set of cognitive mechanisms including restatement, summarization, and judgment/evaluation were used for processing new information with level of abstraction that is moderate compared to generalization and schema induction.

Restatement and summarization have long had their counterparts in natural language processing (NLP) technology: automatic paraphrasing and summarization techniques (Luhn 1958; McKeown, Klavans et al. 1999; Sparck Jones 1999; Nenkova, Passonneau et al. 2007). However, it seemed that by using these cognitive mechanisms, sensemakers were able to make the new information more accessible cognitively in addition to reducing the amount of information required for further analysis. To do so, users often used a language that was easier to understand based on their prior knowledge, and put in perspective in consideration of the task. Having the end products (the paraphrase and the summary) was helpful, but the process of creating such restatement and summary was essential. Because of this, we cannot assume that automatic summarization and paraphrasing alone could substitute for the manual effort put into these processes even if they could yield perfect results.

However, summarization, especially multi-document summarization (McKeown, Klavans et al. 1999) may still assist users when they examine potentially useful articles, because summarization reduces the amount of information, and users can always go back to the original to read in more detail if they find the summary useful. Personalized summarization that takes into consideration the user context might be helpful (Chitrapura, Joshi et al. 2006; Díaz 2007).

Judgment/evaluation was used to form opinions about facts or data found in new information. Such opinions and insights were often expressed in the final product of sensemaking. Users often referred to the opinions and insights by something that "distinguishes" them from other sensemakers. Such opinions may be based on others' opinions in the new information, but more often they were formed through critical thinking and in-depth analysis with a sensemaker's prior knowledge and experience. It seems that sentiment analysis (Pang and Lee 2004; Wilson, Wiebe et al. 2005) may be able to help sensemakers find opinions expressed *by others*, but the process of critical thinking and in-depth reflection cannot be obtained through any automatic means.

Easy annotation and note-taking may be able to assist users by recording any insights that users may come up with during the analytical process. It may also be helpful for the sensemakers to be able to search and browse through the user-generated insights and ideas that come to mind in relation to the acquired facts/data to facilitate the creative part of sensemaking. For example, the relationship between an insight and the note/fact that intrigued that insight may be recorded by the system, and when reviewing the insights the system could show associated notes so that users could easily recall the context where the insight come from.

7.3.1.3 Clustering and Automatic Classification

Clustering creates structure (document clusters) and puts items in each cluster at the same time (Franz, McCarley et al. 2001; Witte and Bergler 2007; Ramage, Heymann et al. 2009) by examining the documents in each cluster users may be able to generate useful structure from these unsorted documents. In this study, users often put documents or results into a general section. Clustering items in this general section may help users in further structure documents that were collected and put into the general section, because they may be potentially useful, but their relationship to the knowledge structure was unclear.

Automatic classification (Hamill and Zamora 2007; Staff and Bugeja 2007; Rajan, Ramalingam et al. 2009) may also help users with manual classification of data items (such as documents) into the conceptual structure (such as note pages). Automatically assigning items to pre-defined categories might have saved users time and effort in manually going through all documents so that they could focus on the important categories.

7.1.3.4 Pattern Recognition and Frame Acquisition

The two mechanisms for processing new information that require most abstraction are generalization and schema induction. Generalization from an example to a class or concept requires domain-specific knowledge. For example, knowing that "Friends" is a TV show may allow a user to generalize from a commercial placed in "Friends" to the general concept of TV ad placement. System assistance may be provided by mapping an example or an instance to an ontology of domain-specific knowledge (Noy, Fergerson et al. 2000; Noy and Musen 2000; Eiter, Ianni et al. 2008) and allowing users to navigate through the knowledge structure. How much it would help the users with identifying concepts through generalization needs further investigation.

The system may assist sensemakers in two ways:

1) Inducing new structure;

2) Filling pre-defined structure with data.

Semantic frame detection (Basili, Croce et al. 2009; Coppola, Gangemi et al. 2009) through lexical patterns may assist users induce new structures. Automatic concept learning (Cohen 2000; Cimiano, Hotho et al. 2005; Lee, Kao et al. 2007) from unstructured text also introduces new structures that may be helpful to sensemakers. How these techniques may used to help with the sensemaking process remains a question to be explored.

Filling pre-defined structure (Cardie 1997; Riloff and Schmelzenbach 1998; Zhou 2007) may assist users with instantiating structures with data. This may be particularly helpful when sensemakers were looking for information about the same facets of multiple targets. For example, when working on finding data to fill in the same frame structure of several ethnic groups as potential markets of a wireless company including size, growth rate, and other facets, User MB6 had to repeat several searches with a set of similar keywords combined with different target groups. The system may be able to save effort for the users by automatically searching for and filling a pre-defined structure with different frames.

7.3.2 Defining and Specifying Concepts

Definition and specification are mechanisms (the only two found in the literature review) used in examining concepts once they are identified. Users sometimes needed definition of a concept as a basis for other operations such as comparison. For example, User MB1 looked up the definition of “stock grant” in order to compare “stock grant” and “stock options”. It would be helpful if such definitions could be automatically acquired from the Web or other sources for the list of concepts that users identify. Users could benefit from automatic identification and extraction of terms along with their definitions from text (Klavans and Muresan 2000; Storrer and Wellinghof 2006).

Having an ontology or dictionary of domain knowledge may help users specify a concept or class with examples; similarly, it may help generalization as discussed in the previous section.

7.3.3 Identifying Relationships and Making them more Salient

Comparison is the fundamental mechanism in the group of mechanisms that were used for examining relationships. Comparison can be used in both bottom-up and top-down fashions to compare two facts or compare to concepts at the structural level. To facilitate comparison of items in the same category, at the minimum a system can display information about two items side by side with highlighting for similarities and differences for easy comparison. For example, users may choose issues of interest and compare candidates' positions on each issue.

If the system is able to detect related concepts or entities (for example, a competing gum company with User MB5's Trident marketing case), the system could help the sensemaker identify relationships between the two by suggesting to the sensemaker to search for several aspects of these similar concepts or entities. The system could even do such a search on its own and present its result in a comparison. It would be useful – if perhaps beyond the state of the art – for the system to induce a schema that applies to multiple items with similar characteristics and then automatically search for data to fill the structure that is pre-defined so that the users may compare different items using the same characteristics or features and identify relationships.

7.3.4 Detection of Anomalies and Inconsistencies

Semantic conflict detection at both the data and schema level (Ram 2004) may help users with examining anomalies and inconsistencies both within the knowledge schema that users constructed and between the knowledge schema and new information. Often a knowledge base is needed for such systems. This ability may help in recognizing surface-level inconsistencies and conflicts in data and schema, but much of the deep and critical thinking that is probed by mechanisms such as Socratic dialog is beyond any automatic means. For example, it would be extremely difficult for any system to ask the critical question *"are they really caring about the environment? Or do they not want to run out of oil and pay a lot for it? <laugh />"* (MJ13) when retrieving the poll result that says people want to develop new energy sources. It would require a lot of background knowledge, experience, and context. That's a creative part of sensemaking that always requires human involvement.

7.4 Assistance at the Sensemaking Task Level

The major goal of many sensemaking tasks is to create an understanding (Stefik, Baldonado et al. 1999), i.e., a structure instantiated with data. Thus an information system aimed to assist sensemaking should provide assistance with structure building and instantiation. Aside from proper representation and manipulation of structure and data in multiple forms, the system should also provide assistance in acquiring structure from other sources, in eliciting structure from users' internal conceptual space, and in linking structure with data.

7.4.1 Assistance with Task Analysis and Gap Identification

It may be helpful to automatically analyze task description and requirements for sensemakers. The minimum function a system may provide is to allow users to enter a description of the task and automatically extract important concepts/entities and relationships in the task description. By examining these concepts and relationships, sensemakers may be able to identify gaps easily.

If the system is designed for a well-defined task type, the system could provide pre-defined task structures and steps for tasks in specific domains as the initial starting point for sensemaking, and allow users to modify the task structures and steps as their sensemaking process continues.

7.4.2 Assistance with Building Structures

The structure elements that compose the sensemakers' conceptual space come from different sources, including:

- Structures from previous knowledge of the sensemaker
- Structure extracted by the sensemaker from task descriptions
- Structure extracted by the system from task descriptions
- Existing structure given in information found
- Structure extracted by the sensemaker from data
- Structure extracted by the system from data

The system should help users' to acquire structures. Structure search and navigation (Qu and Furnas 2007) enable users to search for and browse structures created by others. Structures from multiple sources may be presented to users for comparison, and systems could identify agreement and conflicts between structures from different sources (Ram 2004).

To acquire structure from new information, the system can use techniques discussed in Section 7.4 such as information extraction and schema acquisition to help users identify a preliminary set of entities and relationships to filter. Applying these techniques to task descriptions can also help users to identify structures embedded in the task description. Systems that are designed for specific tasks might provide built-in task templates based on task requirement and workflow analysis.

It is most challenging to elicit structures from the sensemaker's mind. Using multiple representation forms may help users to recognize concepts that they would not be able to detect otherwise. The findings of this research show that by simply changing the representation form (for example, structuring a concept map based on one's notes) without acquiring new information, users were able to elicit additional related concepts and relationships. Users also suggested that being forced to construct the conceptual space in a form that they were not used to sometimes allowed them to think more and recognize structures that could not be seen through their conventional means. The system might also help users come up with their own structures by asking them to answer a list of questions as probes to elicit their existing knowledge.

Once structures are acquired from multiple sources, it is important for the system to differentiate different sources of structure. For example, User MB1 used different colors to differentiate headings directly extracted from case materials and from headings given by the user. The system may allow users to make such distinctions, or, even better, differentiate them automatically when possible. When a structure is from another source, links going back to the original should be kept. When a structure is extracted from data, the data should be included and easily accessible to the user.

7.4.3 Assistance with Instantiating Structures with Data

The system should provide assistance with linking data to structure. There are many ways to do so. For example, in file systems one would have to create a document and save it under a folder in the structure directory. However, this is probably not the best way with many information systems.

The system should make the process of linking data to structure simple and quick so that users do not have to spend a lot of time doing the tedious labor of saving data in its proper place. Such processes do not require much intelligence, but they do require a huge amount of user time. According to (Wright, Schroh et al. 2006) users spend about one third of their time doing reference saving and analysis. They did not specify how time is split between reference saving and analysis, but it is not hard to imagine how time-consuming it would be to create and save files and folders. Moreover, it is also very difficult to access the data once the structures have been instantiated this way. Some means to link data to structure include dragging-and-dropping data onto structure elements, and automatic tracking of sources (URL, file location, and other meta data).

Since it was very common that the processes of building and instantiating structure happen simultaneously or in adjacent pairs (see Section 4.2 for findings on this), it is important that the operations can be done at the same time or right after each other. For example, if a user drags-and-drops data onto a structure element, the system should instantiate that structure element with data; if a user drags-and-drops data onto an empty space in the conceptual space, the system should create a new structure element (for example a concept node) at the same time, and give the user the option of filling in any properties of that concept.

7.4.4 Assistance with Creating Task Products

Most sensemaking tasks produce a task product in some format such as reports, presentations, and plans. Some are more formal than others. These products are largely based on the knowledge structures built during the sensemaking process. Creating these reports usually requires additional work of the users, such as writing a report based on the notes and concept map created. Sometimes the task output may not require much additional effort once the sensemaking is done. For example, a user needed to *"make a review sheet or a sheet to go along with the presentation that I can hand out to the class so they can follow along"* (MB1). In this case, creating the output may be simply selecting important part of a map and particular notes to go with it.

MS OneNote provides a function to save the note pages as a single Word file. This could help users with some tasks such as to prepare notes for discussion (User MB1). But if users need to write a report as the task product, the function does not seem to be useful because most of the notes would not necessarily end up in the written report.

With the knowledge structure accumulated and stored in the system, the system should provide support for creating task outputs out of the knowledge structure in multiple representations. For outputs that require further creative writing, it would be helpful to have a drafting / outlining tool that allows users to select from any representation of the underlying structure (discussed in section 7.1) and *"to take the structure and make an outline, just kind of sorting the different headings*" (MB4). It should also allow users to indicate how much detail they would like to include under the outline headings in terms of which copied notes and annotations they would like to include. This would be especially helpful for writing reports or news articles where users need to refer to some "quotes" or facts to support their claims.

Users also suggested that this kind of mapping tool and note tool would be very helpful when they needed to do immediate reporting and had to file stories online very, very quickly. It would be even more helpful if the mapping tool would fit into a drafting tool that enables users to quickly construct a story.

7.4.5 Assistance with Keeping Track of Sensemaking Process

As noted in the analysis described in Section 4.2, the sensemaking paths that users went through varied from planned and systematic to rather ad-hoc and unsystematic. In many cases, users discovered new leads for searching in the information found (Bates 1989) and went in another direction. Sometimes a search failed temporarily and the user decided to move on and "come back to it later", but never came back to it. As a result, sometimes users lost track of their direction and either got lost or went into a direction that was not likely to get the sensemaker anywhere close to his or her goals (see Section 4.2.3 for details). It might be helpful if users can review their action history (modification of concepts, relationships, and data), and are allowed to go back to a previous conceptual structure in the action history.

Some patterns of the paths users followed suggested a more productive sensemaking process. For example, when users got lost they often went back to the systematic approach, starting with task analysis and gap identification, which allowed them to get back on track.

Users might benefit from automatic monitoring of their sensemaking paths by the system. A system may recognize paths similar to the manually constructed ones described in Section 4.2 and suggest alternative solutions when users seem to be in trouble. For example, if a user has been doing a search for a very long time and is not able to build or instantiate any structure, it is likely that the search is not successful. The system might suggest to the user to use alternative query terms or sources, or it might suggest to the user that she might take a different path, for example, to change the aspect of the topic she worked on, or to look for existing structure rather than try to extract from data. For another example, if a user shows a pattern of repeated "search – instantiating" on a single structure element, the system might suggest that he has collected a lot of information about this particular concept and it is time to work on some concepts that have no data instantiated yet.

It is not an easy task to automatically identify what process a user engaged in. A user may seem to be reading an article for a long time when in fact she is just away from the computer. Different users have different use patterns so that a particular use pattern (for example, browsing) may indicate a search process for one user and a structure building process for another user. Such assistance in recognizing patterns of sensemaking activities has to be customized to individual users, which requires establishing a user profile.

7.5 A Design Framework for Sensemaking Systems / Tools

Table 7-1 describes functions of a system or tool that aims to assist sensemaking. This table is not intended to suggest an exhaustive list of all possible functions, but rather to demonstrate important functions of a sensemaking system and how the functions could support the sensemaking processes (discussed in Section 4.1) and the cognitive mechanisms used in sensemaking (discussed in Section 6.1).

Table 7-1: Functions of the System/Tool with Regard to Sensemaking Processes and Cognitive Mechanisms

Process	Purpose	Function description	Mechanism	Technology
Task analysis and planning	To analyze task and task structure	• Allow users to enter a description of the task • Automatically extract important concepts/entities and relationships in the task description; • Provide pre-defined task structures for tasks in specific domains as the initial starting point for sensemaking	Key item extraction	Information extraction Summarization Knowledge representation
	To identify steps of sensemaking and monitoring processes	• Allow users to define a set of steps to accomplish task goals • For well-defined task, provide pre-defined task steps that the user can edit	Specification	Knowledge representation
		• Show action history (modification of concepts, relationships, and data), and allow users to go back to a previous conceptual structure from history • Monitor the process by identify the search and sensemaking patterns and suggest more productive approaches when users seem in trouble		Action history Discover patterns action history
Gap identification	To facilitate gap identification by task analysis, questioning, and experience	• Allow users to record questions they need to answer • Provide a list of probing questions for well-defined tasks • Detect conflicts at schema and data level	Semantic fit Socratic dialogue Inference	Knowledge representation Expert systems Reasoning systems
	To record different types of gaps and selecting strategies for dealing with them	• Allow users to indicate gaps in the established knowledge structure manually • Differentiate structure and data gaps by color or shape of the concept or relationship in a knowledge structure • Keep track of the status of each gap and provide a status report to users		Visualization Action history
Data seeking and structure seeking	To search for data	• Regular search function that supports Boolean queries, structured queries, and advanced search options		Information retrieval
	To extract concepts and relationships	• To extract entities and relationships from search results, and display in parallel with the original results	Key item extraction	Information extraction
	To search for structure	• The user may select a concept (or multiple concepts) in the work space area, for example concept maps and start a search in various search models • Users could limit search to search for structure elements including entities or relationships. The system should return structure instantiated with its source data.		Information retrieval Knowledge representation Assisted search for structure

Process	Purpose	Function description	Mechanism	Technology
Building Structure	To extract structure from data	• Allow users to select some text in the workspace (such as search results, task descriptions, or reports) and run extraction component on-the-fly • Cluster search results and notes into groups and allow users to create structure based on the clusters • Summarize search results and allow users to build structure based on summaries	Key item extraction Classification Summarization	Information extraction Clustering Summarization
	To support discovery of patterns	• Provide task specific frame structures/templates • Automatically acquire frame structure when possible and fill in structures pre-defined by users • Allow easy comparison of similar concepts and relationships • Allow users to select a piece of text from the search result, and creates a concept and/or a relationship; attach source text is automatically with the notes/arcs created, source information such as URL and date are recorded	Comparison Analogy	
Instantiating Structure	To link data to structure	• Automatic information extraction to fill in the slots of a template or automatic extraction of relationships • Allow users to select a piece of text from the search result and attach it to an existing concept or relationship	Key item extraction Classification	Information extraction Knowledge representation
	To select and organize useful information from the search results	• Information may be instantiated to the structure represented in any form • System provides a set of task-independent tags and task-specific tags • Users can organize the tags into hierarchies or keep loose tags	Classification	Information extraction Summarization
	To differentiate copied notes from annotations	• Allow user to assign tags and notes to search results, concepts, relationships, or attributes in templates • Differentiate copied notes and annotations by automatically added tags • Other visual features to differentiate copied notes and annotations	Judgment / evaluation	
Creating products	To support the outlining and writing of reports	• Allow users to select certain parts of the structure as sub-headings and the outliner creates an outline by selecting appropriate data • Allow users to import concept maps and templates as part of the report		Outlining / editing tools

Process	Purpose	Function description	Mechanism	Technology
Throughout the process: representation and manipulation of conceptual space	To support multiple representations	• The underlying structure can be represented in multiple forms such as concept maps, outlines, text representation, and templates • Automatic transformation from one representation to another		Information representation
	To support users' interaction with the presentation of their conceptual model	• Support representations at different levels of detail of concepts and relationships • Offer the option of differentiating the presentation of different types of concepts and relationships (using different color schemes or highlighting) • Users can zoom in and out to view part of a conceptual model or to get an overview of the whole conceptual space.		Visualization
Throughout the entire process	To keep track of sources	• Copy with source information		Copy with source information

Chapter 8: Conclusions and Future Work

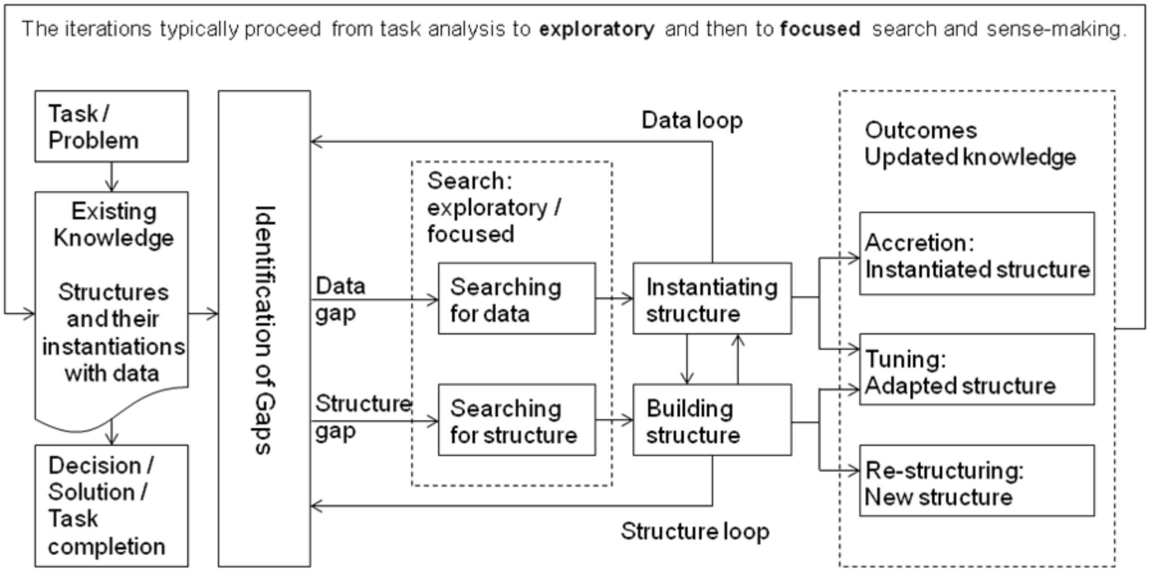
This thesis proposed and examined an iterative sensemaking model (described in Section 2.5) with empirical user studies of 15 users working with business analysis and news writing tasks, and suggested a design framework for information systems and tools that aim to assist sensemaking.

This chapter summarizes the important themes found with the user studies, discusses the theoretical implications and implications for education, and suggests directions for future research.

8.1 Summary of Findings

This section offers a summary of the findings by revisiting some of the important issues raised in the foreshadowing questions and the themes that emerged from the user study, providing an overall picture of the process of sensemaking.

Figure 8-1 illustrates the sensemaking model proposed in this thesis. Sensemaking involves several activities, which may be executed in different sequences depending on the level of existing knowledge and the approach of the sensemaker. The sensemaker's existing knowledge is also iteratively updated in different ways – accretion, tuning, and re-structuring. Several cognitive mechanisms are used in moving the processes along and triggering the conceptual changes. The list of cognitive mechanisms used was augmented with findings of the empirical user study, and the added mechanisms were marked in *italics*.



In each process, the cognitive mechanisms listed below can be used as applicable

Inductive (data-driven, bottom-up)	**Structure-driven (logic-driven, top-down)**	**Other**
Key item extraction	Definition	Comparison
Restatement	Specification	Analogy
Judgment or evaluation	Explanation-based mechanisms	Classification
Summarization	Elimination	*Stereotyping*
Schema induction	Inference	Semantic fit
Generalization		Socratic dialogues

Figure 8-1: Iterative Sensemaking Model (Updated)

(Cognitive mechanisms added)

8.1.1 Iterative Nature of Sensemaking

Several researchers have recognized that sensemaking is not a linear process (Russell, Stefik et al. 1993; Krizan 1999; Stefik, Baldonado et al. 1999; Pirolli and Card 2005), as they model sensemaking as a "cyclical" process and as "sensemaking loops". The findings of this thesis confirmed that the iterative nature of the sensemaking process holds true for business analysis and news writing tasks. This user study also confirms the observation that the simplified waterfall model of how data lead to knowledge and understanding runs counter to empirical evidence (Stefik, Baldonado et al. 1999; Klein, Moon et al. 2006). The model proposed in this thesis views the sensemaking process as composed of several "search – sensemaking" iterations. In each iteration, the sensemaker goes through some search activities (exploratory and focused search for data or structure) followed by some sensemaking activities including gap identification, building structure, instantiating structure and creating products activities (as discussed in Section 4.1).

Findings suggest that the paths that sensemakers went through were rather idiosyncratic and heterogeneous. During each sensemaking iteration, the sensemaker may have selected a different combination of search and sensemaking activities; the manner in which the iterations proceeded varied from one sensemaker to another, depending on the task context, the new information acquired, and the knowledge level of the users.

As idiosyncratic as the paths of search-sensemaking iterations taken by individual sensemakers were, the iterations showed some common patterns that appeared in many cases. The study revealed a spectrum from planned, systematic patterns to rather random, ad hoc patterns of these paths (see Section 4.2 for details). This suggested that the differences lie in *the shift from one iteration to the next* more than within the iterations themselves. This has important implications because such patterns can be captured and analyzed to inform system design and to improve the user's sensemaking approach. Systematic patterns were much less common than ad hoc, unplanned patterns, but they were often more effective in terms of introducing conceptual changes. This suggested the need for careful planning and sensemaking support.

In addition to the iterative process, the knowledge structure created in the sensemaker's conceptual space was also iteratively updated in various ways by accretion, tuning, and restructuring.

8.1.2 Changes in Knowledge Structure and Representation

Changes to the knowledge structure and representation in the external conceptual space reflected the "sense" made during the sensemaking process. The representations were in multiple forms, including text (notes, final report or news story), directory (note folders and pages), and network (concept maps). The most common type of change in knowledge was accretion, which was done through copied notes, restated notes, and summary notes. Accretion by restated notes and summary notes often involved analytical thinking, suggesting the importance of capturing analytical thoughts (Lowrance, Harrison et al. 2001) in sensemaking tasks.

Changes in representation although they did not constitute in themselves modification of either the structure or the data provided a different view of the conceptual space and often led to modifications of the existing schema and introduction of new concepts. For example, the creation of a concept map from note pages often introduced new concepts or made implicit concepts and relationships explicit.

Although related, the degree of change and the type of change are different things. Accretion, tuning, and restructuring are types of changes. Although structure changes (tuning and restructuring) are often more fundamental than the accumulation of facts, an accretion of an important fact may be more significant to the users' knowledge than a less important concept added to the knowledge structure. In the cases under study, structural changes were often gradual. In most cases, users added new concepts gradually as their sensemaking processes progressed. The degree of change was difficult to quantify with measurements. It involved many factors and was often subjective and dependent on the user's existing knowledge.

Prior knowledge plays an important role in sensemaking, similar to meaningful learning (Rumelhart and Norman 1981; Anderson, Reder et al. 1996; Grabowski 1996). In addition to the characteristics of the prior knowledge (Dole and Sinatra 1998), the relationships of the prior knowledge to the to-be-learned knowledge (Chi 2007) also contributed to this role. For some searchers, it was important that new information to relate to prior knowledge, while others could process unrelated information more easily. Information that was complementary to prior knowledge was most helpful in that it added more value to prior knowledge, while information that was in agreement with prior knowledge confirmed what was already known. Information that was in conflict with prior knowledge posed the most challenge to sensemakers; they ended up with acceptance, disregard, partial acceptance of the new evidence, or total confusion.

Tracing conceptual changes to the internal structure was difficult. The think-aloud protocols provided the closest approximation possible. The external representations and internal representations seemed to have different foci and functions in terms of what and how much to externalize and internalize. Users with less domain knowledge seemed to create a larger external representation that focused on the topic structure while users with more domain knowledge seemed to create a smaller external representation that focused on the task structure.

8.1.3 Role of Instantiated Structure

Instantiated structure elements seemed to play an important role in sensemaking:

- Entities (represented as names) and key concepts (represented as keywords) were often the basis for relevance judgments.
- Relationships embedded in new information and between the new information and participants' previous knowledge were crucial in structure building and data fitting.
- Both concepts and relationships seemed to be crucial for updating knowledge. Comparison and dealing with conflicts happened at the level of concepts or overall structure, not at the level of individual data items.

Sometimes building and instantiating structure happened simultaneously when users discovered instantiated structure elements. For example, such instantiated structure may have come from a well structured document where data was put under different headings, or it may have come from a Web site where the site's fits the sensemaker's task so that site structure and the instantiating data could be incorporated. Such instantiated structure elements provided a shortcut for sensemakers, allowing them to take the entire structure along with data. It supported situations when the sensemaker has to build structure but only data was available. Either task of looking for data or building structure was not simple, and was time-consuming at times.

Very often building and instantiating structure happened in successive adjacent pairs. For example, a user may have found a comprehensive document that contains good data, but she has to create the structure from scratch because the document does not provide a good structure that connects with her task. By reading a part of the document, the sensemaker builds a piece of structure and puts some data in it, and then she continues until she finished reading the document. This pattern suggests that the extraction of concepts / entities and relationships (including relationships within the new information and relationships of the new information to existing knowledge) are very important for sensemaking. Users can benefit from more efficient approaches and better system support for successive structure building and instantiation to create instantiated structure elements.

8.1.4 Sensemaking Approaches and the Use of Cognitive Mechanisms

Sensemakers used different ways to approach sensemaking tasks. They used both bottom-up and top-down approaches to sensemaking tasks. Some sensemakers used a combination of bottom-up and top-down approaches. Users' choice of a bottom-up or top-down approach depended on several factors, including the nature of the task (how structured and well-defined the task was), the existing knowledge level of the user (how much the user knew in the domain, the strength, coherence, and commitment of the user's domain knowledge) (Dole and Sinatra 1998), and cognitive styles of the user.

While the overall approach may be bottom-up or top-down, sensemakers used both data-driven and structure-driven mechanisms. The cognitive mechanisms, identified from the literature and complemented by the empirical user study, served several functions in sensemaking. Some were more useful in managing the complexity of the new information and making it more accessible to the sensemaker. Sensemakers also used these cognitive mechanisms to examine the knowledge schema of concepts and relationships and to detect anomalies in knowledge. The mechanisms may have been used alone and in combination with each other. Users' preferences about which mechanisms to use depended on the characteristics of the user and task.

Whether a sensemaker's overall approach is top-down or bottom-up did not seem to affect which cognitive mechanisms he or she used. Users with a primarily top-down approach also used data-driven mechanisms to process new information and connect it with the pre-established structures. Users with a primarily bottom-up approach also used logic-driven mechanisms to examine concepts and infer relationships.

8.1.5 Creativity in Sensemaking

The study confirmed the observation that the process of knowledge building remains a highly creative activity not mastered by automated means (Hsieh and Shipman 2002; Klein, Moon et al. 2006). The creativity may lie in how the sensemaker organized the conceptual space to complete a particular task, or it may lie in the creation of the work product. In this study, the final product of sensemaking was not merely a collection of accumulated knowledge structures and data. It involved the creation of new expressions of what was known and learned, as well as new ideas generated during learning and sensemaking. Such expressions and ideas may have been inspired by structure and data learned when the sensemaker constructed his conceptual space, but they were often original and not found anywhere else in the conceptual space.

In fact, sensemakers sometimes had a strong desire for novelty and deliberately endeavored to achieve it. It did not seem that any special abilities were needed for these creative activities, other than those involved in any other tasks: certain mental abilities such as comprehension and logic, actively open-minded thinking, and expertise in the domain (Baron 1994), which must be acquired over a long period of learning and practice.

*** Sections 8.2 and 8.3 discuss the theoretical and educational implications of the research. Implications for designing information systems and technology that assist sensemaking are discussed in Chapter 7.

8.2 Theoretical Implications

The model proposed in this dissertation provides a descriptive and analytical framework for better understanding of sensemaking processes by extending the existing sensemaking models to theories in cognition and learning. It also sheds light on research in information seeking and use from a perspective of the creation of structured representations. The processes and cognitive mechanisms identified in this thesis provide better foundations for knowledge creation, organization, and sharing practices.

8.2.1 An Analytical and Descriptive Framework for Individual Sensemaking

While the searching aspect has been extensively examined by previous research on sensemaking and by the information retrieval (especially interactive IR) community, research on the construction and use of structured representations has been by and large descriptive. Previous research on sensemaking has focused on the processes and activities that users go through (Russell, Stefik et al. 1993; Krizan 1999; Pirolli and Card 2005; Qu and Furnas 2007), but several important questions about the conceptual changes in the representations sensemakers created and the cognitive mechanisms they use remain unanswered. This thesis provides a framework for analyzing and describing individual sensemaking focusing on the changes to the conceptual space and the cognitive mechanisms used in achieving these changes.

The iterative sensemaking model draws on research in education (especially learning theories), cognitive psychology (cognitive processes and structures), and task-based information seeking and use behaviors. The model provides a stronger basis for explaining sensemaking behaviors by examining users' sensemaking activities, cognition, and the changes to the representations users create while they work on sensemaking tasks with the assistance of sensemaking tools.

By characterizing sensemaking as "search-sensemaking" iterations that are linked with iterative updates of the conceptual space triggered by a set of cognitive mechanisms, the model shows how sensemakers move along from one knowledge state to the next, and what requisites are needed to enable such movements. The common patterns within the iterations and different shift patterns in between iterations suggested that although the sensemaking process as a whole may seem unstructured and idiosyncratic, decomposing the process into a smaller unit (iterations) can show common practices that may be generalizable across different cases and tasks. Although the tasks under study were from specific subject domains, the model provides a comprehensive approach to the integrated design of information systems that incorporates the representation and manipulation of users' conceptual space, information retrieval with respect to structure building, and assistance at both task and cognitive mechanism level.

8.2.2 Task-based Information vs. Structure Seeking and Use

Researchers in library and information sciences (LIS) have been studying task-based information seeking and use, and they made a useful distinction between information task and work task (Vakkari and Hakala 2000; Bystrom 2002). Sensemaking, as an information task, is involved in many work tasks such as problem solving and decision making. The representations constructed during the sensemaking process need to fit the task, or they must be updated (Russell, Stefik et al. 1993). In fact, examining information use as gap-bridging under a sensemaking framework provides insights to information behavior research (Savolainen 2006).

The "search-sensemaking" iterations went from exploratory to focused, to more focused. This is in agreement with findings in task-based information seeking that different types of information (domain knowledge and procedural knowledge) are sought for different types of tasks and/or at different stages of the task. For example, at the beginning (pre-focus) stage, background information is sought, whereas at the end of the task, information that is more specific and pertinent to a chosen focus is used (Kuhlthau 1993; Vakkari and Hakala 2000). The seeking for meaning (Kuhlthau 2004) should include both structure seeking and fact seeking. Sometimes structure seeking may be more dominant than fact seeking.

Task-based information seeking and use research can benefit from the analysis of creation and use of structured representations for tasks and problems. In addition to the compound nature of information tasks and work tasks, the examination of the concepts and relationships in the knowledge space of users suggested that task structure and topic structure are often intertwined and work together to best serve the functional demands of the task.

The cognitive mechanisms identified in this research also shed light on task-based information seeking and use, in that they may be able to provide some explanatory power about the cognitive drivers for triggering information seeking behaviors. Reasons for starting and ending sensemaking iterations may also shed lights on information seeking models such as the berry-picking model (Bates 1989).

8.2.3 Knowledge Creation, Organization, and Sharing

Research in knowledge management has been concerned mainly with the creation, transfer and sharing of knowledge in an organizational setting (Alavi and Leidner 2001). Research in knowledge creation has mainly focused on the social and cultural aspects of organization.

According to Nonaka (1994), knowledge creation has four modes:

- Internalization
- Externalization

- Socialization
- Combination

Essentially, internalization (the creation of new tacit knowledge from explicit knowledge by learning and understanding) and combination (the creation of new explicit knowledge by merging , categorizing, reclassifying and synthesizing existing explicit knowledge) are very similar to the sensemaking process that users went through to create their internal and external conceptual space (internal and external) from multiple sources. The model proposed in this thesis may shed light on individual knowledge creation by identifying the processes that knowledge workers go through and the different ways the knowledge structure is modified. Sensemaking tools that facilitate the elicitation of knowledge structures from the internal space may also be useful in the externalization mode in knowledge creation.

The sharing of knowledge should be concerned not only with knowledge as objects, but also with knowledge as process. Sharing not only the end product of individual sensemaking but also the intermediate structures and processes for creating such structures may be helpful. Research in knowledge sharing technology has identified the importance of heterogeneous representations of knowledge (Neches, Fikes et al. 1991). As Stefik et al. (Stefik, Baldonado et al. 1999) research on knowledge sharing technology can benefit from design ideas for sensemaking tools, for example, as discussed in Chapter 7.

8.3 Implications for Education

It was brought up by participants in the exit interviews that participating in the research session had an impact on their ways of approaching sensemaking tasks in general, and they suggested that sensemaking tools and skills should be introduced early in high school or even in middle school. Research in information literacy education has focused on the ability to recognize an information need and to locate, evaluate, organize and use information in various settings such as learning, problem-solving, and decision-making in formal and informal learning contexts (American Library Association 1989; Bruce 1997; Eisenberg 2008). While a lot of emphasis has been put on the ability to "find information", less attention has been paid to skills in making sense of the information found, and to the use of sensemaking tools. An important competency for sensemaking is the ability to represent, convey, and acquire structural knowledge using concept maps and other visual organizers (Jonassen, Beissner et al. 1993; Hyerle 1996).

8.3.1 Sensemaking Skills Education

Although some researchers claim that no evidence for a general sensemaking skill has been seen, other researchers have identified general cognitive skills that are useful to sensemaking, such as the skill of noticing anomalies and inferring causal mental models, and the ability to induce rich schemas (Klein, Moon et al. 2006).

Research in teaching general cognitive skills (Perkins 1985) suggested that such skills can be acquired through proper instruction, and an individual's capacity for acquiring and using information can be enhanced by training in appropriate information-processing strategies. The education of sensemaking skills and strategies may benefit from research in learning strategies. Dansereau (1985) describes a complex learning strategy system which is composed of primary strategies and supportive strategies.

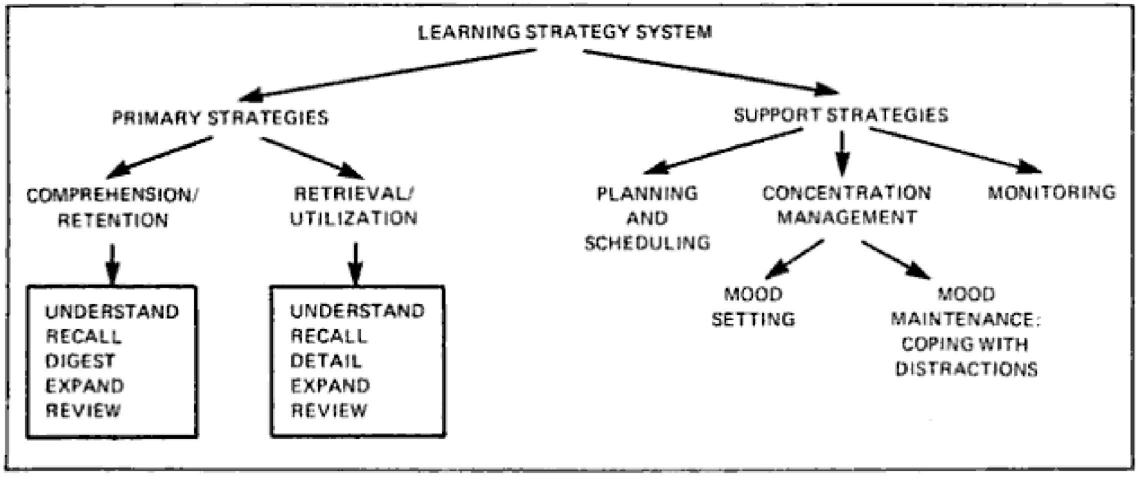


Figure 8-2: Overview of the Learning Strategy System

Dansereau (1985), p.219

Similar to the learning strategy system, sensemaking strategies and skills can also be divided into primary and supportive categories. The primary skills for sensemaking have to do with the ability to recognize gaps, to interpret information, to create structure from data, and to fit data into structure. The supporting skills for sensemaking have to do with task planning, monitoring, and concentration management to keep track of the sensemaking process. The following figure shows a sensemaking skills system that is composed of primary and supportive skills:

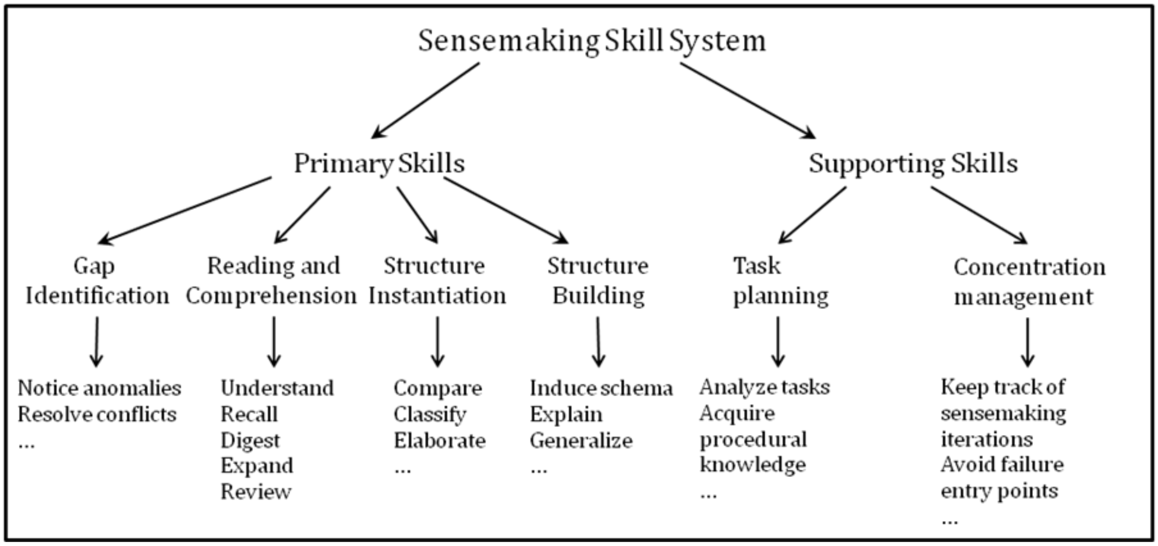


Figure 8-3: Primary and Supporting Sensemaking Skills

This figure is just a starting point for sensemaking skills education. Some cognitive mechanisms used in sensemaking may translate to cognitive skills, and others may be acquired through general education on reading, writing, and logic. Further research and analysis are needed to identify a list of skills and capabilities that are related to sensemaking, especially to structure building and instantiation.

Sensemaking skills and habits of approaching tasks should be formed early in education; these skills have to be introduced early in the students' learning, preferably in secondary and middle schools. In terms of pedagogical approach and design training materials, research on content-independent learning strategies suggested that there is a tradeoff between the specificity of a learning strategy or skill and the likelihood of training students to use that skill (Dansereau 1985). This is likely to hold true for sensemaking skills. On the one hand, training scenarios must be designed to be general enough to detach from one particular sensemaking case to focus on the skills; on the other hand, they must be similar to tasks that students work with to provide opportunities to apply general skills to specific problems.

8.3.2 Introduction of Sensemaking Tools

Users suggested that it would be useful to introduce tools that assist sensemaking, such as the tools used in this study, in a class and to teach students how to use them early in their education because habits of intellectual work form early. Users were able to think of other task scenarios in which they would use the tools after using them in the user study session.

Similar to the education of general and context-specific sensemaking skills, some sensemaking tools are specific to certain tasks, such as argumentation tools (Cho and Jonassen 2002; Gersh, Lewis et al. 2006) and others are more generic, such as concept mapping and note-taking. Generic tools cannot be introduced without any task context, but the task should be not too specific to restrain the users from learning to use the tools for other tasks.

More importantly, rather than teaching the functionalities of specific software applications, students should be taught about investigating and selecting tools for their tasks. They should have the basic awareness that such a category of tools is available, and should be able to select the most appropriate tool for a particular task. This is not to claim that all students should be able to make design suggestions for information systems and technology. Rather, students should have the knowledge about what could be done by systems and what technologies may be available to help them achieve their tasks.

8.4 Future Directions

Researchers have sometimes avoided talking about internal representations and cognitive aspects of sensemaking because of the difficulty in assessing what is in a user's mind (Cacioppo and Petty 1981; Das 1994; Chi 2006) and the limitations of using verbal reports and observations (Nisbett and Wilson 1977; Ericsson and Simon 1993; Hoffman, Shadbolt et al. 1995) as data to interpret mental process. However, the cognitive processes and mechanisms are fundamental to information behavior research and need more attention in the field.

As an initial inquiry into the conceptual changes to users' knowledge structure and cognitive mechanisms used in sensemaking, this study has uncovered a rich set of findings related to the creation and use of structured representations in sensemaking. These findings provoke more questions in regards to the knowledge structure in sensemaking. These questions include:

1. What roles do topic structure and task structure play in sensemaking? How are topic structure and task structure combined? Where do the concepts and relationships come from and how are they connected?
2. How should the degree of conceptual changes be measured?
3. What is the relationship between internal structure and external structure? How do they work together for a sensemaker? What are their functions and roles? How can representational and other aids help users to make explicit their internal conceptual structures?
4. Can this model be transferred to a collaborative setting to describe and analyze collaborative creation of structured representations? What other issues need to be accounted for in collaborative settings?

5. Can external knowledge structures such as ontologies help users acquire concepts and relationships? How?

There are more questions waiting for answers than the few listed here. Further analysis of the data collected in this study may shed light on some of these questions. As the next step, a detailed concept analysis of the structured representations that users created in this study may be able to shed some light on **topic vs. task structures** without collecting additional data. Some possible measurable factors that have an effect on **the degree of a conceptual change** include the concept's placement in the knowledge structure (for example, its depth in a knowledge hierarchy and its centrality in a network), the contribution of a fact (whether it introduces new concepts or just confirms existing ones, and how important the concepts are), and whether they appear in the final work product. This needs further investigation.

Sensemaking in collaborative and organizational settings. The model proposed in this paper deals with individual sensemaking, but much sensemaking activity occurs in groups. It would be interesting to see whether the iterative model of sensemaking can be extended to collaborative learning (educational) and computer supported collaborative work (CSCW) (organizational) settings without much change to the basic framework. Collaborative sensemaking also involves conceptual changes to the collective knowledge structure and to each collaborator's knowledge structure. They may be analyzed using the same framework. In addition to the cognitive mechanisms each individual member used, there may be mechanisms and activities for facilitating communication among members. Future work should include examining conceptual changes and mechanisms in collaborative settings and investigating how system tools may assist collaborative sensemaking. Tools that enable systematic note-taking and well-structured external representation of knowledge schemas may be even more important for people working together in sensemaking.

Design and evaluation of tools. This thesis has provided a design framework for information systems and technologies to assist sensemaking. It seems a good idea to implement some design ideas starting with a particular task domain, such as business analysis, news writing, or informal learning. Any sensemaking support system should incorporate the core idea of having an underlying architecture with multiple representations, focus on facilitating structure acquisition, and provide supports at the task and cognitive mechanisms levels. How users use such tools needs more investigation.

Sensemaking education. The first step would be to identify a set of skills that are essential to sensemaking to amend the skill system shown in Figure 8-3. This would require intensive examination of literature in areas such as cognitive and thinking skills education, composition and writing education, and general literacy education. Also needed is investigation of the relationship between cognitive mechanisms used in sensemaking and cognitive skills that people can be trained to use. The next steps would be designing training scenarios and materials for selected user groups and conducting studies to evaluate the effectiveness of such training. Guidelines may be produced on how to educate people to become better sensemakers based on findings from these studies.

8.5 Significance of the Study

1. **A new sensemaking model.** The main contribution is the iterative sensemaking model proposed and examined, building upon previous sensemaking research, learning theories, cognitive psychology and task-based information seeking and use. The model provides a descriptive and analytical framework in examining sensemaking process which may be applied to settings beyond the user cases examined in this research.
2. **Sequence diagram for visualizing and analyzing the sensemaking process and activities**.
3. **Recurring modules of the sensemaking process.** Research in sensemaking has discovered the idiosyncratic and iterative nature of sensemaking; this dissertation reveals the recurring modules of the sensemaking process – various ways in which the sensemaking iterations started and ended. Findings suggested that the heterogeneous patterns of sensemaking lie in the shifts from one iteration to the next, rather than in the iterations themselves.
4. **Understanding of conceptual changes and cognitive mechanisms.** Guided by the model, the results lead to a better understanding of conceptual changes that occur in the sensemakers' knowledge structure and the cognitive mechanisms used to trigger these changes during the sensemaking process. These results deepen the

understanding of information behavior, task-oriented learning, and eventually organizational learning.

5. **Recommendations for the design of sensemaking support systems.** The design framework and recommendations for design functionality based on the case studies provide the basis for improved sensemaking support systems.

Beyond information retrieval, supporting users in making sense of the information found is the next frontier in information research. The major contribution of this work is the framework and model in linking the iterative sensemaking processes and activities with the conceptual changes in knowledge structure and cognitive mechanisms that underlie the activities and conceptual changes. The model and findings in the empirical user study provided a better basis for system design.

Appendix A: Definitions

Sensemaking: The process "by which individuals (or organizations) create an understanding so that they can act in a principled and informed manner" (Stefik, et al., 1999). Researchers used different spellings for sensemaking: Sense-Making (Dervin, 1980, 1992, 1998), sensemaking (Qu & Furnas, 2007; Russell, Stefik, Pirolli, & Card, 1993; Stefik, et al., 1999), or sense-making (Romano, Bauer, Chen, & Nunamaker, 2000; Savolainen, 2006).

In the literature and in this paper, the term "sensemaking" has both a broad and narrow meaning. The broad meaning, for example (Pirolli & Card, 2005), refers to the total process of (1) searching for information that is relevant for a task and (2) through a process of further extraction, analysis, and integration creating an understanding on which to base decisions or actions. The narrow meaning is restricted to (2), the processes of relating information found to previous knowledge, creating structures and fitting data into structures to create representations, and thus arriving at an understanding of a situation or phenomenon (Russell, et al., 1993). It is generally clear from the context which meaning is intended. Sensemaking models are about sensemaking in the broad meaning.

Sensemaker: Te agent of the sensemaking activities. A sensemaker could be an individual, a group of individuals, an organization, or possibly even a computer program. This dissertation focuses on individual sensemakers.

Knowledge: There are several definitions of knowledge from philosophy, epistemology, and education. The sensemaking research takes the constructivist view and defines knowledge as "product of and fodder for sensemaking and sense unmaking" (Dervin, 1998). Knowledge is the sense made by someone at some time. In this dissertation, knowledge is understood as the structure instantiated by data that the user constructs in the process of sensemaking. It may be represented with less structure such as natural language expressions or with formal structures such as network representations.

Representation: In this thesis, the term "representation" includes both structure and data organized in a meaningful way (structure instantiated with data, or instantiated structure). Representations are reflections of users' knowledge of a particular task or problem. Representations may consist of structural elements (entities, concepts, and/or relationships among entities and concepts) and data that support them. In the literature, the term "representation" is sometimes used to mean just structure.

Structure: a fundamental notion underlying patterns and relationships of entities. In this thesis, structure is used as a general term to encompass patterns, schemas, frames, and other terms with similar meaning.

Structure elements: the components that constitute a structure. A structure element can be an entity, a concept, or a relationship and may be represented in various formats.

Structural knowledge: structural knowledge is generally defined as the knowledge about the structure of concepts in a knowledge domain and can be measured in a variety of ways (Jonassen, Beissner, & Yacci, 1993). Structural knowledge may be captured with network representations, but networks cannot function adequately as the sole means of representation.

Consuming knowledge structure: the process of using structure instantiated with data for accomplishing a task (Russell, et al., 1993). The consumption of knowledge structure is closely related to comprehension, which is often reduced to the use of some knowledge structures for other cognitive activities such as inference (Graesser & Clark, 1985).

Concept: the term *concept* is a loaded one within cognitive psychology, philosophy, linguistics, and related disciplines. Often concept is taken to mean a mental representation of a simple class. In the scope of this research, the focus is put on the function of concepts as cognitive states for sensemaking, with emphasis more on a concept's relationships to other concepts, and less on its categorization functions.

Concept maps: graphical tools for organizing and representing knowledge (Novak, 2006). They include concepts, usually enclosed in circles or boxes of some type, and relationships between concepts indicated by a connecting line linking two concepts.

Task: In this thesis, the unqualified term *task* refers to *work task* (Kim & Soergel, 2005; Vakkari & Hakala, 2000) as opposed to *information task*. Information tasks are used to include both search tasks and sensemaking tasks. For example, to develop a marketing plan for a company is a work task, whereas to search for and to make sense of information in order to develop the plan is information tasks.

Appendix B: Data Collection Instruments

B.1 IRB Approval Forms

See next three pages.

UNIVERSITY OF
MARYLAND
INSTITUTIONAL REVIEW BOARD

2100 Blair Lee Building
College Park, Maryland 20742-5121
301.405.4212 TEL 301.314.1475 FAX
irb@deans.umd.edu
www.umresearch.umd.edu/IRB

December 14, 2007

MEMORANDUM

Application Approval Notification

To:	Dr. Dagobert Soergel Pengyi Zhang College of Information Studies
From:	Roslyn Edson, M.S., CIP IRB Manager University of Maryland, College Park
Re:	**IRB Application Number: # 07-0672** **Project Title:** "Supporting Sense – making with Tools for Structuring a Conceptual Space"
Approval Date:	**December 13, 2007**
Expiration Date:	**December 13, 2008**
Type of Application:	New Project
Type of Research:	Non-Exempt
Type of Review For Application:	Expedited

The University of Maryland, College Park Institutional Review Board (IRB) approved your IRB application. The research was approved in accordance with 45 CFR 46, the Federal Policy for the Protection of Human Subjects, and the University's IRB policies and procedures. Please reference the above-cited IRB application number in any future communications with our office regarding this research.

Recruitment/Consent: For research requiring written informed consent, the IRB-approved and stamped informed consent document is enclosed. The IRB approval expiration date has been stamped on the informed consent document. Please keep copies of the consent forms used for this research for three years after the completion of the research.

Continuing Review: If you intend to continue to collect data from human subjects or to analyze private, identifiable data collected from human subjects, after the expiration date for this approval (indicated above), you must submit a renewal application to the IRB Office at least 30 days before the approval expiration date.

Modifications: Any changes to the approved protocol must be approved by the IRB before the change is implemented, except when a change is necessary to eliminate apparent immediate hazards to the subjects. If you would like to modify the approved protocol, please submit an addendum request to the IRB Office. The instructions for submitting a request are posted on the IRB web site at: http://www.umresearch.umd.edu/IRB/irb_Addendum%20Protocol.htm.

UNIVERSITY OF MARYLAND
INSTITUTIONAL REVIEW BOARD

2100 Lee Building
College Park, Maryland 20742-5121
301.405.4212 TEL 301.314.1475 FAX
irb@deans.umd.edu
www.umresearch.umd.edu/IRB

July 16, 2008

MEMORANDUM

Application Approval Notification

To:	Dr. Dagobert Soergel Dr. John Newhagen Pengyi Zhang College of Information Studies
From:	Roslyn Edson, M.S., CIP IRB Manager University of Maryland, College Park
Re:	**IRB Application Number: # 07-0672** **Project Title:** "Supporting Sense-making with Tools for Structuring a Conceptual Space"
Approval Date:	**July 16, 2008**
Expiration Date:	**December 13, 2008**
Application Type:	***Addendum / Modification;*** Approval of request, submitted to the IRB office on July 15, 2008, to (1) add Dr. John Newhagen as a Co-Investigator (2) add an interview which will be conducted at the end of the semester when the research is conducted (3) analyze data collected from the perspective of journalism education (4) collect additional data (5) add a consent script for the additional data collection
Type of Review of Addendum:	**Expedited**
Type of Research:	Non-Exempt

The University of Maryland, College Park Institutional Review Board (IRB) approved your IRB application. The research was approved in accordance with 45 CFR 46, the Federal Policy for the Protection of Human Subjects, and the University's IRB policies and procedures. Please reference the above-cited IRB application number in any future communications with our office regarding this research.

Recruitment/Consent: For research requiring written informed consent, the IRB-approved and stamped informed consent document is enclosed. The IRB approval expiration date has been stamped on the informed consent document. Please keep copies of the consent forms used for this research for three years after the completion of the research.

Continuing Review: If you intend to continue to collect data from human subjects or to analyze private, identifiable data collected from human subjects, after the expiration date for this approval (indicated above), you must submit a renewal application to the IRB Office at least 30 days before the approval expiration date.

UNIVERSITY OF MARYLAND

INSTITUTIONAL REVIEW BOARD

2100 Lee Building
College Park, Maryland 20742-5121
301.405.4212 TEL 301.314.1475 FAX
irb@deans.umd.edu
www.umresearch.umd.edu/IRB

November 20, 2008

MEMORANDUM

Application Approval Notification

To: Dr. Dagobert Soergel
Dr. John Newhagen
Pengyi Zhang
College of Information Studies

From: Mary Ann Ottinger, Ph.D. MAO
Associate Vice President for Research Compliance and Policy
University of Maryland, College Park

Re: **IRB Application Number:** 07-0672
Project Title: "Supporting Sense-making with Tools for Structuring a Conceptual Space"

Approval Date: November 20, 2008

Expiration Date: December 13, 2009

Type of Application: Renewal

Type of Research: Non-Exempt

Type of Review for Application: Expedited

The University of Maryland, College Park Institutional Review Board (IRB) approved your IRB application. The research was approved in accordance with the University IRB policies and procedures and 45 CFR 46, the Federal Policy for the Protection of Human Subjects. Please include the above-cited IRB application number in any future communications with our office regarding this research.

Recruitment/Consent: For research requiring written informed consent, the IRB-approved and stamped informed consent document is enclosed. The expiration date for IRB approval has been stamped on the informed consent document. Please keep copies of the consent forms used for this research for three years after the completion of the research.

Continuing Review: If you intend to continue to collect data from human subjects or to analyze private, identifiable data collected from human subjects, after the expiration date for this approval (indicated above), you must submit a renewal application to the IRB Office at least 30 days before the approval expiration date. If IRB approval of your project expires, all human subject research activities including the enrollment of new subjects, data collection, and analysis of identifiable private information must stop until the renewal application is approved by the IRB.

Modifications: Any changes to the approved protocol must be approved by the IRB before the change is implemented, except when a change is necessary to eliminate apparent immediate hazards to the subjects. If you

B.2 Recruitment Flyer

Looking for participants
For a study in computer-assisted sense-making

Overview	**What:** I am looking for about 10 participants for a research project, which involves using innovative tools for analytical tasks. Students would participate in: • a 1-hour training session (to prepare you for using the tools), and • a 2- or 3-hour assignment session (to work on an assignment from the course). The tools include MS OneNote (part of MS Office 2007) for note-taking, free concept mapping software for organizing thoughts and analysis, and Word for the final writing. (Other tools such as Excel may also be used.) Participants will use a two-screen work station equipped with custom installation of the software. **Who:** Students in BMGT440 Advanced Financial Management **When:** Fall 2008 (Sep 3 – Nov 6) **Where:** 4111B Hornbake Building (South Wing), Univ. of Maryland
Benefits	You will learn about cool analysis and writing tools. You will be paid $45 for the 3-hour participation ($15 for the training session and $30 for the first assignment you do). If you do a second assignment, you will be paid an additional $30.
How to participate?	Please fill the sign-up sheet if you are interested in participating and return it to me at the end of the class. Please sign up for the possible date and time when it is convenient for you (can be adjusted later by email). At the training session you will be asked to sign an informed consent form. All students are invited to a demo session (30 minutes) introducing the tools [date/time/location] or [date/time/location]. Any questions or did not have a chance to sign up? Please email me.
Contact	Pengyi Zhang PhD Candidate, College of Information Studies Email: pengyi@umd.edu Tel: 240-481-4224 More about the project can be found at www.wam.umd.edu/~pengyi/sensemaking

Sign-up Sheet

Name: ______________________________

E-mail: ______________________________

Phone: ______________________________

Please put down possible dates and time for the training session and assignment session(s). Make sure you sign up for a date/ time which will leave you enough time to complete your assignment before it is due.

Training Session (during Sep 10 – Sep 24)

Assignment Session(s) (during Sep 22 – Nov 3)

	Training session (1-hour) Date / Time	Individual case assignment (2-hour) (Due on Nov 3) Date / Time
Option 1		
Option 2		
Option 3		

B.3 Consent Form

Page 1 of 2
Initials _______ Date ______

CONSENT FORM

Project Title	Supporting Sensemaking with Tools for Structuring a Conceptual Space
Why is this research being done?	This is a research project being conducted by Pengyi Zhang and Dagobert Soergel in the College of Information Studies at the University of Maryland, College Park. We are inviting you to participate in this research project because you are working on tasks that require you to understand a topic or situation. The purpose of this research project is to investigate how users make sense of task situations using the assistance of an information system, and to design better tools to facilitate such processes.
What will I be asked to do?	The procedures involve three task sessions. In each task session, you will participate in: 1. **Training.** You will be given a short introduction of the system. Training will be given only on the first task session. 2. **Evaluation tasks.** You will complete a task (assigned to you, or a task of your own), in which you will use the system for searching for information and organizing your understanding. During the task, you will be asked to think aloud – verbalizing your thoughts related to the task. 3. **Feedback.** Before and/or after each task, you may be asked to provide feedback in the form of questionnaires or interviews. The think-aloud protocols and feedback interviews will be audio-taped. Sessions will meet on the University of Maryland Campus. We hope you can attend all three sessions.
What about confidentiality?	We will do our best to keep your personal information confidential. To help protect your confidentiality (1) your name will not be included on the surveys and other collected data; (2) a code will be placed on the survey and other collected data; (3) only the researcher will have access to the identification key. All electronic records of your work will be maintained on a secure disk separate from your identifying information. This research project involves making audiotapes of your think-aloud protocols during the task and feedback interviews after the task for analysis of the sensemaking processes. Only researchers of the project will have access to the data which will be destroyed three years after the completion of the project. Quotations from the interviews and think-aloud protocols may be used in the research reports. ___ I agree to be audiotaped during my participation in this study. ___ I do not agree to be audiotaped during my participation in this study. If we write a report or article about this research project, your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.

Project Title	Supporting Sensemaking with Tools for Structuring a Conceptual Space
What are the risks of this research?	There are no known risks associated with participation in the research project.
What are the benefits of this research?	Benefits to participants: participants will learn strategies for organizing information from disparate sources which may be useful for their future work; Benefits of the research: the research will advance our understanding of how people make sense of information for their tasks, and thus better tools may be developed to help users of information systems make sense of the information they find.
Do I have to be in this research? May I stop participating at any time?	Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify.
What if I have questions?	This research is being conducted by **Prof. Dagobert Soergel** and Pengyi Zhang at the University of Maryland, College Park. If you have any questions about the research study itself, please contact: Dagobert Soergel, College of Information Studies 4105 Hornbake Building, Univ. of Maryland, College Park, MD, 20742 (email) dsoergel@umd.edu (telephone) 301-405-2037 If you have questions about your rights as a research subject or wish to report a research-related injury, please contact: **Institutional Review Board Office, Univ. of Maryland, College Park, Maryland, 20742; (e-mail) irb@deans.umd.edu; (telephone) 301-405-0678** This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects.
Statement of Age of Subject and Consent	Your signature indicates that: • you are at least 18 years of age; • the research has been explained to you; • your questions have been fully answered; and • you freely and voluntarily choose to participate in this research project.

Signature and Date	**NAME OF SUBJECT**	
	SIGNATURE OF SUBJECT	
	DATE	

B.4 User Background Questionnaire

Part 1: Basic Information

1. User Code: ________________________

2. Gender:

☐ Female ☐ Male

3. Age: ________________________

4. Which year? (Check one)

☐ Freshman ☐ Junior
☐ Sophomore ☐ Senior
☐ Other, please specify ________________________

5. Major: ________________________

6. Which of the following courses are you enrolled in?

☐ BMGT 440 ☐ LBSC 635 ☐ JOUR 320
☐ BMGT 450 ☐ INFM 612 ☐ JOUR 471
☐ BMGT 452 ☐ Other, ______

7. What languages do you speak?

Native language(s): ____

Fluent in: ____

Part 2: Computer Use

1. On average, how much time do you spend per day using a computer? (Check one)

☐ Less than 1 hr ☐ 4 hrs. – less than 5 hrs.
☐ 1 hr – less than 2 hrs. ☐ 5 hrs. – less than 6 hrs.
☐ 2 hrs. – less than 3 hrs. ☐ More than 6 hrs.
☐ 3 hrs. – less than 4 hrs.

2. Which Web browser(s) do you most often use?

☐ Safari ☐ Netscape
☐ Internet Explorer ☐ Other, please specify ________
☐ Mozilla Firefox

3. What operating system(s) do you most often use?

☐ Windows Vista ☐ Mac Leopard
☐ Windows XP ☐ Mac Tiger
☐ Windows 2000 ☐ Other, please specify ________

4. Which of the following search engines or systems do you often use? (Select all that apply)

- ☐ Google
- ☐ Yahoo! Search
- ☐ AOL Search
- ☐ MSN Search
- ☐ AltaVista
- ☐ Ask.com
- ☐ Dogpile
- ☐ Excite
- ☐ Infoseek
- ☐ Inktomi
- ☐ Lycos
- ☐ WebCrawler
- ☐ Baltimore Sun
- ☐ Bloomberg
- ☐ CNN.com
- ☐ Factiva.com
- ☐ Lexis/Nexis
- ☐ New York Times Online
- ☐ Reuters.com
- ☐ The Economist Online
- ☐ The Financial Times Online
- ☐ Wall Street Journal Online
- ☐ Washington Post Online
- ☐ USAToday.com
- ☐ Other, please specify____

Part 3: Problem Solving

1. What computer program do you usually use to do your assignments?

- ☐ MS Word
- ☐ Word Perfect
- ☐ MS Spreadsheet
- ☐ MS Powerpoint
- ☐ Concept mapping software
- ☐ Other, please specify ___________

2. How confident are you in your abilities to locate specific information using a search engine? (Circle one)

1 2 3 4 5 6 7

Not at all confident Very Confident

3. Have you ever used concept mapping techniques, i.e., drawing concepts and relationships of a topic or situation?

- ☐ Yes, please describe ____

  ____

  ____

- ☐ No.

4. Have you used any of the following software?

- ☐ Mind / Idea / Concept mapping software, such as CMapTool, Mind Meister, Mind42, IMindMap, Tinder Box or Personal Brain
- ☐ Note-taking software, such as MS OneNote, Word Perfect Lightening, TinderBox, AMNotes, or Golden Section Notes
- ☐ None of the above

B.5 Training Objectives

Tool Training Objectives Check List

OneNote

- ☐ Understand how the note pages are organized
- ☐ Know how to add, edit, or delete a note
- ☐ Know how to organize notes visually
- ☐ Know how to search notes
- ☐ Know how to save notes

CMapTool

- ☐ Know how to add, edit, and delete nodes
- ☐ Know how to add, edit, and delete relationships
- ☐ Know how to organize a map, use colors and shapes to show patterns
- ☐ Know how to link a map to Source URLs and/or Notes in MS OneNote

B.6 Training Material

Note: The training materials are similar in content but the examples are different for the business cases and the journalism cases. This example is for the business cases.

Part One: Instruction of the tools

Training Scenario: Your task is to search the Internet, learn about the current banking crisis with Fannie Mae and Freddie Mac, and prepare an analysis of US government's taking control of the mortgage giants.

The general work flow involves the following phases:

5. Searching for information using Firefox or IE;
6. Taking and organizing notes in MS OneNote 2007;
7. Organizing your thoughts for your case analysis in CMap;
8. Prepare your analysis in Word 2007.

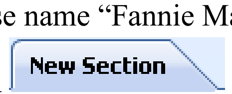
The shortcuts for these programs are listed on the taskbar next to the "Start" menu. 

0. Using two screens

You can move your mouse freely between two screens as if you were working with one screen.

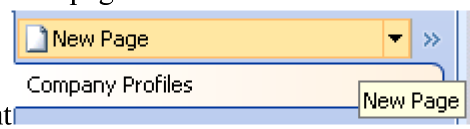
If you would like to move an application to another screen, you need to do so when the application window is NOT maximized.

1: Planning

1. Start Microsoft Office OneNote 2007
2. Create a new notebook named with your user code. To do this, click on File → New → Notebook, and enter your user code as the name of the notebook. Use the default setting for other options.
3. OneNote creates a new section and a new page.
4. Name the new section with your case name "Fannie Mae and Freddie Mac". To do this, right-click on the tab named 

and select "Rename". Type in "Fannie Mae and Freddie Mac".

5. Name the current page “Company Profiles”. To do this, type in the title in the blank box on the current page:

6. Create another new page titled “Problem Definition”: click on “New Page” button on the top right 

, and type in “Problem Definition”.

7. Since you would like to evaluate alternative solutions to the problem, develop a plan for the best alternative, and, evaluate the solution, create three pages for “Alternative Solutions”, “Implementing Plan” and “Evaluation”.

8. You may create other pages for other information to be included in your analysis.

9. On the right page bar, you can reorder the pages by clicking on one page, and drag it to the desired position. For example, put “Problem Definition” before “Company Profiles”.

Tip: Notes are organized in a “notebook → section → page → (subpage →) note” structure. A note is usually some text with formatting, but a note can also be a picture, or a table. Notes may be freely moved and arranged in a page.

2: Collecting information and taking notes in OneNote

Taking notes with printed materials

1. Read the first two paragraphs of the Economist article “The muddle-through approach”. Take a few notes in your “problem definition” page to summarize the situation. To do so, click on anywhere that you would like to start with, and start typing in your notes. You can move the note boxes around to arrange them.

2. You can change the format of the notes to ruled lines (just like a paper-based notebook) . To do so, click on “Format → Rule lines” and select the format you like.

Taking notes from Web material

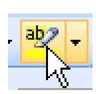
3. Start Firefox or IE, do a search on “Fannie Mae” in the search engine or database you normally use, for example, www.google.com
4. Find some information about the company that you think will be useful in your analysis. For example, go to its Wikipedia page, and find some basic profile information about the company.
5. Copy and paste the information into appropriate pages you just created in OneNote.

Tip: It is much easier to work with IE/Firefox on one screen and OneNote on the other screen. You can highlight a paragraph in the IE/Firefox window and drag it into OneNote.

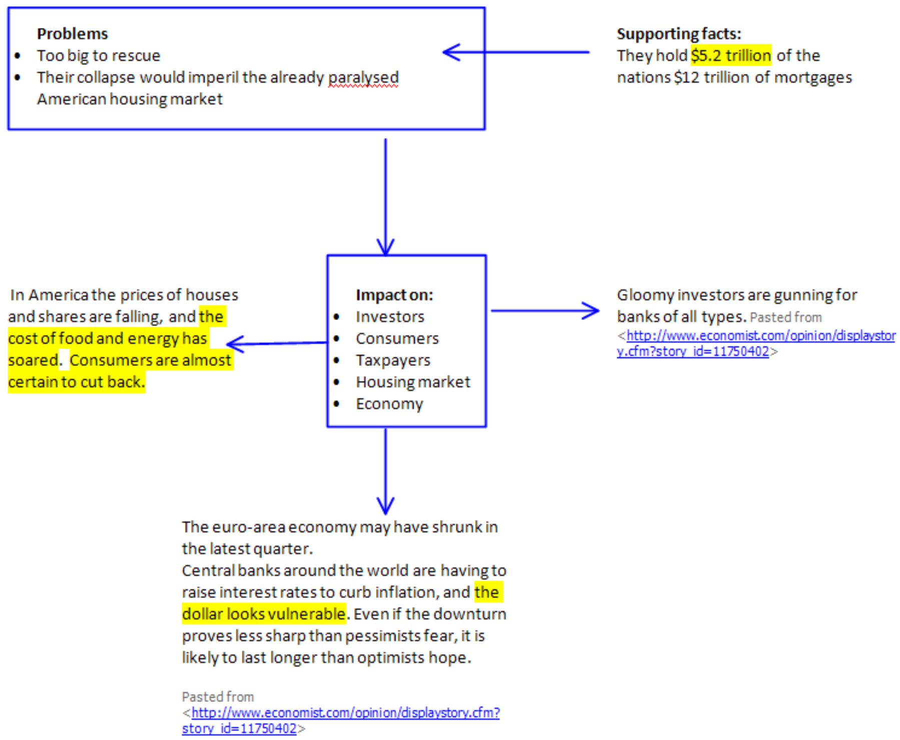
6. Note that OneNote has created a note with the source URL of the website automatically attached.
7. You can edit or resize the note if you need to.
8. Find some information on “Freddie Mac” and take notes on that too.

Tip: Images can be copied and pasted into OneNote as a note.

3. Organize notes visually

1. Go back to the Problem Definition page in OneNote
2. Click on the border of a note.
3. Move the note to another place in this page. To do so, left-click the mouse, the note will be highlighted. Drag the note box to the desired place, and release the mouse.
4. Since you can have several notes on one page, it is a good practice to organize them visually in a meaningful way.
5. You can use the highlighting function  to highlight some of your notes.
6. You can draw boxes around notes and draw arrows among them using the drawing tool bar. The drawing tool bar is located at the bottom of the OneNote window. To access the drawing tool bar, click on “View → Drawing Toolbar”

Tip: An example of several notes arranged in a page is shown below:



4. Save notes

Note are saved automatically as MS OneNote files. Notes can also be saved as other formats include MS Word by clicking on "File → Save as…" and selecting the desired format.

1. Save the notes as a Word document under My Documents\Training\, name it with your user code;
2. Open the Word file to see how the notes are organized in the file.

5. Structure your analysis in CMap

CMap Tool is used to convey the overall visualized analysis of the information collected and notes taken. You may extract the concept map from search results or your notes taken in MS OneNote. A concept map is called a CMap in the CMapTool.

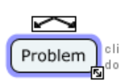
1. Click the CMapTool icon . This opens a "Views" window showing all your saved folders and CMaps.
2. If you see a window named "untitled", go to Step 3. Otherwise, create a new CMap by clicking "File → New CMap" (or pressing Ctrl+N). This opens a new untitled CMap window.

3. Save this CMap as "MortgageCrisisMap": in this newly opened CMap window, click "File → Save CMap" (or pressing Ctrl+S) and enter the name of the map "MortgageCrisisMap".
4. MortgageCrisisMap is saved under root folder of the program. You can see it in the "Views" window under the icon 
.
5. Now create a folder named with your user code: click "File –> New Folder" (or pressing Ctrl + Shift + N) and enter the folder name. This folder will also appear in the "Views" window under the icon 
.
6. Drag and move MortgageCrisisMap into the training folder.

Tip: Concept maps are organized in folders similar to the Windows directories. You can double click a folder name or a CMap name to open a folder or a CMap.

Create, edit and delete a node

In CMap, a map contains two types of elements: nodes and links. You can use a node to represent a concept, an entity, or an event.

1. Go back to the MortgageCrisisMap window. If it is closed, double click "MortgageCrisisMap" shown in the "Views Window".
2. Double click at the place where you would like to create a node (there maybe some delay). This creates a node labeled with "????".
3. Click on the "????". This allows you to type in your own label for the node. Enter "Problem", for example.
4. Create some other nodes such as "Freddie Mae", "Government", "Impact", "Consumers", "Investors", "Economy", and "Alternative Solutions".
5. Note that if you click on a node, the border of this node is highlighted 
, and you can resize the node by dragging the little icon on the right-bottom corner of the node.
6. Creates some other nodes that you think are important to include
7. Save the CMap (Ctrl + S).

Tip: you can right-click on the node to cut, copy, paste or delete the node. You can use "Ctrl + Z" to undo an operation.

Create, edit, and delete a relationship

1. Click on the node "Problem" to have it highlighted.

2. Drag the icon directly into the node "Impact" to connect to it. This creates a link labeled "????" between the two nodes. Double click on the "????" and enter your own label for the relationship, such as "causes".

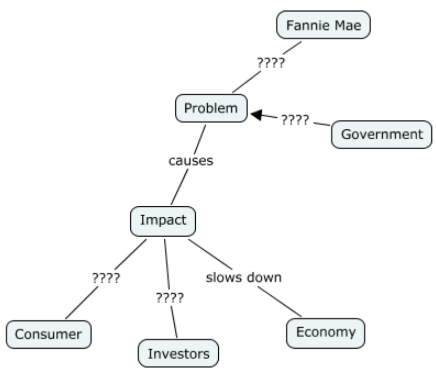
3. Create another relationship to indicate that "Impact" has different dimensions on "Consumers", "Investors", and "Economy".

4. Drag the icon above a node into an empty space. This creates a relationship and a new node at the same time.

5. Create other meaningful relationships (and nodes) as to how you would analyze the situation.

6. Save the CMap (Ctrl + S).

Below is an example map. Your map for a real task/assignment may include more information and be organized differently such as by timeline, causal relationships, etc.



Attach Source URLs and/or Notes in MS OneNote to a node

Attach a Source URL

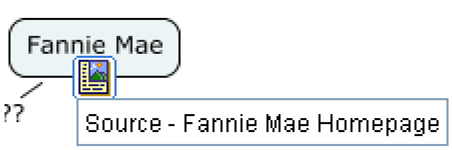
1. Find the homepage of Fannie Mae, and copy the URL in the address bar.

2. Go back to the MortgageCrisisMap map, right-click on the "Freddie Mae" node and select "Add Web Addresses".

3. Fill in the following:
   Resource Name: "Source – Fannie Mae Homepage"
   Resource Type: "URL"

4. Click inside the Web address box, and press Ctrl + V to paste the address.

5. Press "OK". A small icon will appear under the "career path and achievement" node:



6. Click on the icon, you will see a label "Source – Fannie Mae Homepage" which is the name of the resource that you just entered.
7. Click on the label. This will open the company's homepage in a Firefox (or IE) window.

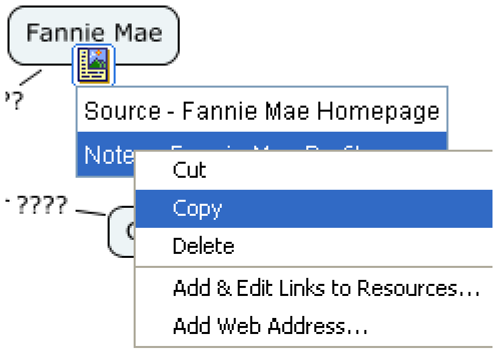
Attach a Note from MS OneNote

1. Go back to your OneNote application.
2. On the "Company Profile" page, highlight the note that has Fannie Mae's information.
3. Right-click on this note, and select "Copy Hyperlink to this Paragraph".
4. Go back to the MortgageCrisisMap.
5. Right-click on the "Fannie Mae" node and select "Add Web Addresses".
6. Fill in the following:
   Resource Name: "Notes – Fannie Mae Profile"
   Resource Type: "URL"
7. In the Web address box, paste the hyperlink you just copied from OneNote application (Ctrl + V).
8. Press "OK".
9. Click on the icon under "Fannie Mae", and then click on "Notes – Fannie Mae Profile". This will go to the note you attached to this node. If OneNote is not open, it will automatically open the OneNote application and go to the note that you attached.

Copy a resource link from one node to another node

1. Click the icon under "Fannie Mae" node, you should see all the links you added earlier.

2. Right-click on "Notes – Fannie Mea Profile", select "Copy".



3. Right-click on the node "Problem", select "Paste".

4. A small icon appears under the "Problem" node. Click on it to see if the link has been copied.

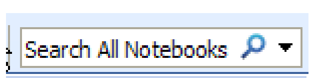
5. Click on the label to go to the source or go to the notes you have taken.

6. Creates a few more links to your notes and the websites you found.

Tip: By attaching source URLs and notes to the nodes in a concept map, you can organize all the information you collected about a concept (or entity) in one place. The hyperlinks make it easier to navigate through different programs when you prepare your analysis based on the maps and notes.

6. Search Notes

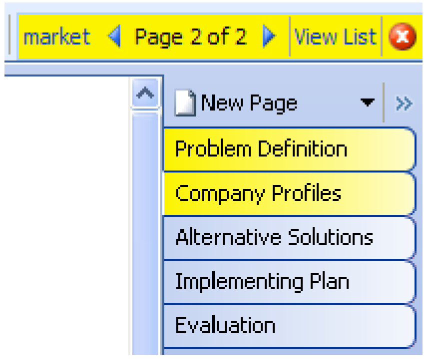
Suppose you are now looking for all notes that mention "market". To find all the notes you have that are related to "market", you can search your own notes using keyword search.

1. Click inside the 

box just above the "New Page" button on the top right.

2. Type in your keyword to search the notes, such as "market", press "Enter".

3. The notes in which your keyword is present are highlighted, and you will see a information bar like this



4. Press the  arrow to view the next match. This will go through all the notes that have the keyword. You can attach these notes to the "market" Node in your CMap.
5. If multiple pages contain the keyword, you can see a list of all the pages by clicking on "View List" button.
6. To exit this search, press "X".

Tip: You can search her notes using keyword search. This is especially useful when you would like to get all the notes about a keyword when you are trying to attach links to your CMap nodes and when you are at the writing stage to compile all related notes you have taken about a particular issue.

7. Write your analysis

The map you created in CMap is a perfect outline or structure to start your analysis. From the map, the links you added allows instant access to the sources and to your notes in OneNote.

You can play with these tools with your own computer.

MS OneNote is part of the Office 2007 package.

CMapTools is free for personal use. To download CMapTools, go to http://cmap.ihmc.us/download/

Part Two: Think-aloud Exercise

Think aloud instruction

While working on your assignment (or the assigned task), you are asked to "think-aloud", that is, to explicitly say aloud whatever is going through your mind. Please keep the following in mind:

1) Verbalize your thoughts as much as possible. Try to avoid silent periods.
2) Focus on your task, and merely verbalize your thoughts; do not talk as if you are trying to explain it to anyone else.
3) Don't worry about the grammar or completeness of your utterance. Although articulating your thoughts clearly is preferred, you should not feel burdened in doing so.
4) Say aloud what you are thinking about, not what you are doing. For example, if you are reading an article and interesting ideas come to your mind, for example, it confirms or contradicts with what you know about the issue, say aloud such ideas, not "I am reading an article".

Please continue to work on the Freddie Mac and Fannie Mae Case using the tools and practice thinking aloud as you go.

B.7 Assignment Session Instruction

1. Do your own research using search engines that locate polling results as well as scholarly and journalistic sources, and write a 400-word story about the role of energy, including surrounding factors such as global warming, as an issue in the election. The story can be an overview of the issue, or you can focus the topic to a specific facet of the issue.

2. Use **OneNote** for note-taking, **CMapTool** for structuring your article, and **Word** for writing the article. You have up to 2 hours and 40 minutes to work on the story. Spend enough time gathering your notes and structuring your map, and also leave enough time to write your story.

3. While you are doing the assignment, remember to **think aloud – verbalizing ALL the thoughts that go through your mind.**

   1) Speak clearly. Don't worry about grammar or completeness of your sentences.

   2) Avoid silent periods.

   3) Say aloud what you are thinking, not what you are doing.

4. At the end of the session, you will be able to save and send your notes, maps, and the story to yourself. The instructor will announce the deadline to turn in the story in class.

B.8 Post-Session Interview Questions

1. What is your general experience with this assignment session?
2. What did you know about xyz (topic of the task)?
3. Please describe the whole processes you went through to make the notes and map. (What's your approach to collect information and organize the story/analysis?)
4. What did you expect to find when you started?
5. Let's look at your notes. How did you organize your notes? How did you come up with the labels of the pages (if any)?
6. How did you decide what to put into your notes when you find something useful? (ask for examples)
7. Let's go to CMap, could you explain your map to me?
8. How did you created your nodes and labeled them from the detailed notes? (then ask for examples)
9. Do you have any criteria as to how you would label/organize the nodes and their links?
10. How did you (or how are you going to if the user has not done writing) write the news article / answer case questions / create the case presentation, based on the notes and the maps?
11. What do you think are the most important things you learned about the issue/case?
12. Is there anything you were looking for but did not find?
13. Do you think you might use these tools for other tasks of your own?
14. Do you have any suggestions to make the tools work better to assist your task?

B.9 Exit Interview Questions

Sensemaking approach:

1. How do you compare the experience of writing story with the tools with what you usually do for stories?
2. How did the experience of using the tools affect the way you your approach writing (or case) assignments or other sensemaking tasks?
3. What are the most important aspects of your sensemaking task and how do you approach them with/without computer tools?

Tools:

1. Did you start using the tools on your own? If yes, what tasks do you use them for? How do you use them for these tasks? If not, why?
2. What do you think the tools would be most useful for?
3. When you worked on the assignment with the tools, what are the most useful features to you? What do you like most about the tools?
4. What do you dislike about the tools? Do you have any suggestions to improve the tools that you used? What functions did you wish to have?
5. Do you have any suggestions for more integrative tools? What would an ideal tool look like? How could the current tools be improved to be more helpful to you?

B.10 Member Check Questions

Please read the file **Appendix E: Trident Marketing Case, User MB5**, and rate from 1 to 5 on the following questions as best as you can recall.

Not at all	Not quite	Sort of	Moderately Accurate	Very Accurate	Do not recall at all
1	**2**	**3**	**4**	**5**	**0**

1. Does the description of the task accurately summarize your task when you participated in the study?
2. With the think aloud table E-2:
    a. Does column B user activity accurately describe your steps?
    b. Does column C think-aloud protocol accurately reflected what you said at the time?
    c. Does the highlighted coding make sense to you? (Can you understand what they are with the textural explanations below?)
3. Does the "brief description" in Table E-3 accurately describe what you went through working on this task in general?
4. Does the note structure shown in Table E-4 accurately reflect what you did with OneNote?
5. Does the concept map shown in Figure E-1 accurately reflect what you did with CMap?

If rated with a 4 or below, please give an explanation of what was missing or incorrect so I can improve it.

Appendix C: Participant Characteristics

Part I: Background					
User Code	**Gender**	**Age**	**Year**	**Major**	**Language(s)**
mb1	M	21	Senior	Finance	English (N)
mb3	F	24	Graduate, 2nd year	Information management	Chinese (N) English (F)
mb4	F	21	Senior	Finance	English (N)
mb5	M	20	Junior	Marketing	English (N)
mb6	F	21	Senior	Marketing	English (N)
mj1	F	19	Junior	Journalism English	English (N)
mj3	F	21	Senior	Journalism Government and Politics	English (N)
mj4	F	21	Senior	Journalism American Studies	English (N)
mj5	F	21	Senior	Journalism Government	English (N)
mj8	M	22	Senior	Journalism	English (N) Spanish (F)
mj9	M	21	Senior	Government and Politics Journalism	English (N)
mj10	F	22	Senior	Journalism	English (N)
mj13	F	21	Senior	Broadcast journalism	English (N)
mj14	F	21	Senior	Print Journalism	English (N)
mj15	F	21	Senior	Journalism Business management	English (N)

Part II: Computer use				
User Code	**Computer use (hours per day)**	**Web browser**	**Operating system**	**Search engines or systems**
mb1	2-3	Firefox	Windows XP	Google, New York Times Online
mb3	5-6	IE	Windows XP	Google, Baidu
mb4	2-3	Firefox	Windows XP	Google, CNN.com, Washington Post Online
mb5	3-4	Firefox	Windows XP	Google, New York Times Online, Washington Post Online
mb6	4-5	IE, Firefox, Chrome	Windows XP	Google
mj1	3-4	Firefox	Windows XP	Google, Lexis/Nexis, New York Times Online, Washington Post Online
mj3	3-4	Firefox	Windows Vista	Google, CNN.com, New York Times Online, Washington Post Online
mj4	5-6	Firefox	Windows 97	Google, New York Time Online, Washington Post Online
mj5	more than 6	IE	Windows Vista	Google, Yahoo Search, CNN.com, Washington Post Online
mj8	2-3	Firefox	Mac OS X	Google, Yahoo! Search, AOL Search, CNN.com, Lexis/Nexis, New York Times Online, Washington Post Online
mj9	2-3	SaFari	Windows XP	Google, Lexis/Nexis
mj10	2-3	IE Firefox	Windows XP Mac Leopard	Google, Washington Post Online
mj13	5-6	IE	Windows XP	Google, AOL Search, CNN.com, Lexis/Nexis, Washington Post Online
mj14	more than 6	SaFari IE Firefox	Windows Vista Windows 2000	Google, Lexis/Nexis, Washington Post Online
mj15	3-4	Firefox	Windows XP	Google, Yahoo! Search, Lexis/Nexis, Washington Post Online

Part III: Problem Solving				
User Code	**Software used now**	**Software used before**	**Familiar with concept mapping as an idea?**	**Confidence in searching ability (1-7)**
mb1	MS Word	Note-taking	N	6
mb3	MS Word MS Spreadsheet	Concept mapping	Y, to draw concepts and relationships of a research topic	7
mb4	MS Word MS Spreadsheet	None	Y, outlines For paper-thoughts and organizing main points	6
mb5	MS Word	None	N	6
mb6	MS Word MS Spreadsheet MS PowerPoint	None	Y, sequential diagrams For organizing papers	6
mj1	MS Word	None	N	5
mj3	MS Word	None	N	5
mj4	MS Word MS Spreadsheet	None	N	5
mj5	MS Word	None	No	5
mj8	MS Word	None	N	6
mj9	MS Word	None	Y, I used concept mapping to organize a research paper I wrote this semester	5
mj10	MS Word	None	N	5
mj13	MS Word	None	Y, I use webs/lists before I write longer stories For TV/the web.	6
mj14	MS Word	None	Y, in class, particularly say 465, we've been using diagrams to show relationships b/w concepts	6
mj15	MS Word	None	Y, briefly used concept mapping for 2 classes I had at UMD; we used it to brainstorm and show relationships between ideas in a project paper.	5

Appendix D: Transcription Guideline and Conventions (V3)

This guideline applies to transcribing think-aloud protocols and user activities as recorded by the Camtasia software.

D.1 Format

The transcripts should include 3 columns: 1. time stamps, 2. user activities, and 3. think-aloud protocols. An example transcript looks like this:

Time	User activity	Think-aloud protocol
00:00:17:: 00:00:56	**B2** Started Firefox, went to Google Web, and did a Google search with keywords "energy election 2008"	**C2** Okay [00:00:18::00:00:30] Right now I'm just going to do a Google search. Okay. This is on <sp /> "the role of energy, including surrounding factors such as global warming as an issue in the election." Okay I'm just going to Google "energy election 2008".
00:00:57:: 00:01:03	**B3** Clicked on the first result, titled "NPR 2008 Election Issues Climate Change"	**C3** I should look <sp /> "NPR 2008 Election Issues Climate Change".
00:01:04:: 00:01:17	**B4** Reading the first paragraph and browsed the rest of the website http://www.npr.org/news/specials/election2008/issues/climate.html Copied and pasted the overall positions of the two candidates into OneNote.	**C4** This is a nice graphic and a nice, um, overview of what both candidates want, so I am going to paste these things into OneNote. Okay, so…
00:01:18:: 00:02:40	**B5** Reading unintelligibly the notes copied into OneNote about both McCain and Obama's positions on climate change.	**C5** <x>reading unintelligibly<x/>
00:02:41:: 00:03:02	**B6** Went back to the website, copied another paragraph into the OneNote page.	**C6** Okay.
00:03:03:: 00:03:30	**B7** Attempted to give a title to the OneNote page, but did not come up with a title	**C7** I'm going to title this… What the title is? <sp /> Climate change… Um…
00:03:31:: 00:03:37	**B8** went back to The NPR website, browsed the hyperlinks, and clicked on a link to an article titled "Interest in Climate Change Heats Up in 2008 Race" http://www.npr.org/templates/story/story.php?storyId=10840816	**C8** connections … interesting climate change heats up.
00:03:38:: 00:04:25	**B9** Browsed the article, copied and pasted the whole article into OneNote	**C9** Oh, this looks like a good article [00:03:42::00:04:02] I'm just going to copy this whole article and put it in because it all looks good.

Time is transcribed to the seconds, for example: "00:03:38::00:04:25" indicates that the time span is from 00:03:38 to 00:04:25. Time spans should be continuous in adjacent rows. For example, the next row should start at 00:04:02. Use [] for long pauses inside think-aloud protocols, for example [00:03:42::00:04:02]. For short pauses, use <sp />.

The timing should be based primarily on user activity. If users' think-aloud started or ended a bit earlier or later than the activity, use the activity time as the time for the row, but put the think-aloud protocol into the cell which it logically belongs to.

Use natural breaks to decide when to start a new row. For example, usually when a user starts a new activity, or switches to a different application, a new row should be started. Sometimes if an activity takes very long, for example, reading an article in Web Browser or OneNote, or drawing a map in CMap, it should be divided into smaller units.

The cells should be labeled as B2, B3, B4… for the activity column and C2, C3, C4… for the think-aloud column.

D.2 Think-aloud Transcription Conventions

Meaning	**Annotation**	**Note**
Pause		**Pauses that are 1 second or shorter are not transcribed. Pauses from 2-4 seconds are considered short pauses; and pauses that are 5 seconds or longer are considered long pauses.**
Short Pause	<sp/>	Short pauses are marked as </sp>
Long Pause	[00:03:22::00:04:01]	Long pauses are timed using [start time::end time] format. Start time and end time are formatted in minutes and seconds as HH:MM:SS.
Break	<break> [00:03:22::00:12:11] description </break>	Breaks for pauses unrelated to the assignment or task, such as technical problems, researcher checking in, Q&A with the researcher, etc. Provide a short description of what the break is about.
Tone		
Mild emphasis	 	
Strong emphasis	 	
Miscellaneous		
No activity screen grab worth noting	No Activity	Indicates that no notable activity took place in this time interval.
Reading aloud	“”	Use quotes for reading
Call-out	In writing you can’t start a sentence with the word ‘and’.	
Grammatical mistake	[sic]	Follows mistake to indicate accuracy of transcription
Cut-off thought	Which is hysterical, because … So what I want to do next.	Three dots indicate the cut off thought
Unintelligible	<x/>	Use <x/> for unintelligible disfluency without approximation
Uncertain hearing	<x> He is going. </x>	Use <x> </x> for disfluency with approximation for uncertain hearings with the hearings in between.
Word truncation/cut-off	Wor-	
“um", "uh", or "mm"	um, uh, or mm	

Meaning	**Annotation**	**Note**
Laughter	<laugh />	
Coughing	<cough />	
Yawn	<yawn />	
Sneeze	<sneeze />	
Sigh	<sigh />	
Other effects	• <ironic> </ironic> • <sarcastic> </sarcastic> • <incredulous> </incredulous> • <frustrated> </frustrated> • <surprise> </surprise> • <satisfaction> </satisfaction> • <dissatisfaction> </dissatisfaction> • <certainty> </certainty> • <uncertainty> </uncertainty> • <parenthetical> </parenthetical>	Use these codes only when plain transcription does not clearly reveal the emotional content of the utterance.
Observer comments	**<oc> comments </oc>**	**Use <oc> </oc> for observations that the transcriber thinks is important to make a note on.**

D.3 User activities

Types of user activities to be transcribed (including but not limited to the following list):

1. Applications
    - 1.1 Starting a new application. For example, user started Internet Explorer
    - 1.2 Switching from one application to another. For example, user moved from Internet Explorer to OneNote.
    - 1.3 Closing an application
2. OneNote:
    - 2.1 Creating a page
    - 2.2 Giving a title to a page
    - 2.3 Creating a note (box)
    - 2.4 Editing a note (adding to a note, modifying or deleting part of a note)
        - 2.4.1 Copying notes from a website
        - 2.4.2 Taking notes manually
    - 2.5 Formatting a note (box)
    - 2.6 Moving a note (box)
    - 2.7 Deleting a note (box)
    - 2.8 Highlighting notes
3. CMap:
    - 3.1 Creating a node
    - 3.2 Labeling a node
    - 3.3 Modifying a node
    - 3.4 Moving a node
    - 3.5 Changing format of a node
    - 3.6 Deleting a node
    - 3.7 Creating a link
    - 3.8 Labeling a link
    - 3.9 Modifying a link
    - 3.10 Moving a link
    - 3.11 Changing format of a link
    - 3.12 Deleting a link
4. Other:

4.1 Searching using a keyword

4.2 Browsing a website

4.3 Reading an article/ a paragraph in IE or Firefox

4.4 Typing in Word

Format of transcripts:

1. Use MM:SS::MM:SS to mark the starting and ending times of the activity, followed by a description of the activity.
2. Labels of pages, notes, nodes or links, phrases/words from the notes should be in quotes and in *italic.*

Appendix E: Trident Marketing Case, User MB5

This appendix presents a detailed case in the study. It includes five parts:

E.1: a description of the case abridged from the assignment package provided by the user.

E.2: an example think-aloud excerpt with coding and explanations of the coding.

E.3: a case description that was derived from the coding illustrated in E.2

E.4: note pages and concept map created by the user.

E.5: the sequence chart created from the code.

E.1 Assignment Description (Abridged Version)

Trident Integrated Marketing Communications (IMC) Project You are developing an integrated marketing communications (IMC) plan for a gum product, Trident, including a TV advertisement and two other advertising and promotion media such as print advertising, radio advertising, billboard advertising, direct marketing, web marketing, telemarketing, direct sales, consumer or trade promotions, etc. Gather current advertisements from your product and its competitors, conduct thorough research in trade and business periodicals on the product, the company, competitors, category users, category trends, and market shares. Develop multiple ideas for the plan based on the research you have conducted. Your plan should address the problem of the company and the objectives of your proposal, analyze the current marketing and advertising situation, recommend IMC strategy and tactics, and discuss alternatives that you considered but rejected.

Table E-1: Assignment Description (Trident Marketing Case)

E.2 Think-aloud Protocol with Coding, MB5

Time	B. User activity	C. Think-aloud Protocol	Processes	Conceptual Changes	Cognitive Mechanisms
…	B13 …	C13 …	..	…	…
09:20:: 09:53	B14 created a new page in OneNote "Problem/Opportunity"	C14 I just want… okay. [0918::0929] problem and opportunity… and I just need something creative to try to separate us from the rest of the completion. May be why people try gum might be helpful.	Building structure Concept "problem and opportunity" recognized Gap identification Identified data gap for reasons why people chew gum	Tuning Added a concept "Problem/O pportunity" to the structure	
09:54:: 10:47	B15 continued to read the article in Firefox copied and pasted a paragraph about adults and teens being more likely to use regular than sugarless mints into OneNote under "Problem/Opportunity" page.	C15 Right here it says most people like… "prefer the sugarless gum", well, "but adults and teens are more likely to use regular gum", so that could be part of the differentiation just to advertise to really point out the sugarless fact and try to sell that. Since it does not seem like people really prefer the sugar. [SP]	Focused search for data Looked for reasons why people chew gum Instantiating structure Linked facts to "problem/ opportunity" concept	Accretion Added a fact	Key item extraction Items extracted: people prefer the sugarless gum Comparison Compared the adult and teen audiences
10:48:: 11:25	B16 copied another paragraph into the "Problem/Opportunity" page; continued to browse the article.	C16 "Children and teens are more likely to chew gum". So that can be something where if we want to either target the children or try to capture the adult market. It says "among users, adults chew 8 pieces and teens 11 pieces". [SP] this is too much information [11:10::11:20]	Focused search for data Looked for reasons why people chew gum Instantiating structure Linked facts to "problem/ opportunity" concept	Accretion Added some facts	Key item extraction Extracted items:" Children and teens are more likely to chew gum", "among users, adults chew 8 pieces and teens 11 pieces".

Time	B. User activity	C. Think-aloud Protocol	Processes	Conceptual Changes	Cognitive Mechanisms
11:26:: 12:49	B17 copied and pasted a paragraph into OneNote "Problem/Opportunity" page, and continued to browse the article	C17 "the adults are less likely than teens to use gum and breath mints", so I think that's something. [11:27::11:35] I want a table that shows me why people are chewing gum. This information is very helpful, like, who chew gums and demographics, but I want to know why. Demographics are helpful when we are doing advertising itself but now I am trying to get what special feature we want to advertise for our gum, so I really need something that says why people are chewing. This is just demographic information. [SP] I don't want this… "Trident flavors", that's interesting [SP] Oh here we go. No… [SP]	Focused search for data Looked for reasons why people chew gum Instantiating structure Linked facts to "problem/ opportunity" concept Gap identification Needs data for reasons why people chew gum	Accretion Added some facts	Key item extraction Extracted items: "the adults are less likely than teens to use gum and breath mints", "Trident flavors"
12:50:: 13:18	B18 copied and pasted another paragraph into "P/O" page in OneNote, continued to browse	C18 "Rules and etiquette", starting to get at habit of chewing and why people chew and things like that. [12:47::12:06]	Focused search for data Looked for reasons why people chew gum Instantiating structure Linked facts to "problem/ opportunity" concept	Accretion Added some facts	Key item extraction Extracted item: "Rules and etiquette"
13:19:: 14:47	B19 copied two bullets from the section of "Interest in functional gum"	C19 Oh this is what I need. All right perfect. "Using gum as a delivery system"… I mean again this is helpful. Tables are just too much for me to look at right now. This is perfect information, okay. [SP] okay so this is starting to get some reasons people are chewing gum. The types of gum people like.	Instantiating structure Linked facts to concept "functional gum"	Accretion Added some facts	Key item extraction Extracted item: "using gum as a delivery system"
…	B20	C20	…	…	…

Table E-2: Example Think-aloud Protocol with Coding, User MB5

E.3 Case Description of Search-Sensemaking Iterations, MB5

	Brief Description	Paths		Conceptual changes	Cognitive mechanisms
		Search	Sensemaking		
1	Analyzed task requirement, did a general search on "gum", found basic knowledge about gum sector, put into his notes	Exploratory search	Building structure Instantiating structure	Accretion Tuning	Specification Key item extraction
2	Created a "problem/opportunity" page, found information about that and put into the notes; looked for reasons "why people chew gum" and found some information and put it into her notes	Focused search for data	Building structure Gap identification Instantiating structure	Accretion Tuning	Key item extraction Comparison
3	Created a map in CMap, added new concepts "functionality", "adult market", "young audience", "whitening recipe", and "new flavors", added relationships		Building Structure Updating knowledge	Tuning	Comparison Semantic fit Socratic dialogue
4	Searched for Trident marketing, found some paragraphs talking about what Trident needs to do, put it into OneNote	Focused search for data	Instantiating structure	Accretion	Key item extraction Generalization Semantic fit
5	Tried to search in a database from the library, did not want to register to log in.	Focused search for data, failed			
6	Searched for trident advertisements, found only ads from UK, did not find much TV advertising	Focused search for data, failed			Key item extraction
7	Browsed a report found earlier, looked for ideas for innovation, come up with some ideas for advertising, added the concept "needs" to the concept map	Focused search	Building structure Instantiating structure	Accretion Tuning	Key item extraction Comparison Analogy and metaphor Schema induction Generalization Elimination
8	Searched for more information on "Trident", found some information and added the concept "ad ideas"	Focused search for data	Gap identification Structure building	Tuning	Key item extraction Inference
9	Started writing the report, searched for specific percentage of market share (failed) and the name of an intergradient (successful); searched for previous ads of Trident and ads from Orbits, come up with concrete ideas for the campaign.	Focused search for data	Gap identification Instantiating structure	Accretion	Key item extraction Comparison

Table E-3: Search-Sensemaking Iterations, User MB5

E.4 Note Pages and Concept Map (MB5)

Trident
Gum Sector
Competitors
Orbit
Advertisements
Problem/Opportunity
Adult/traditional
Youth
Innovation
Ad Ideas
Two ideas
Previous ads of Trident
Ads from Orbit

Table E-4: Note Pages and Structure, User MB5

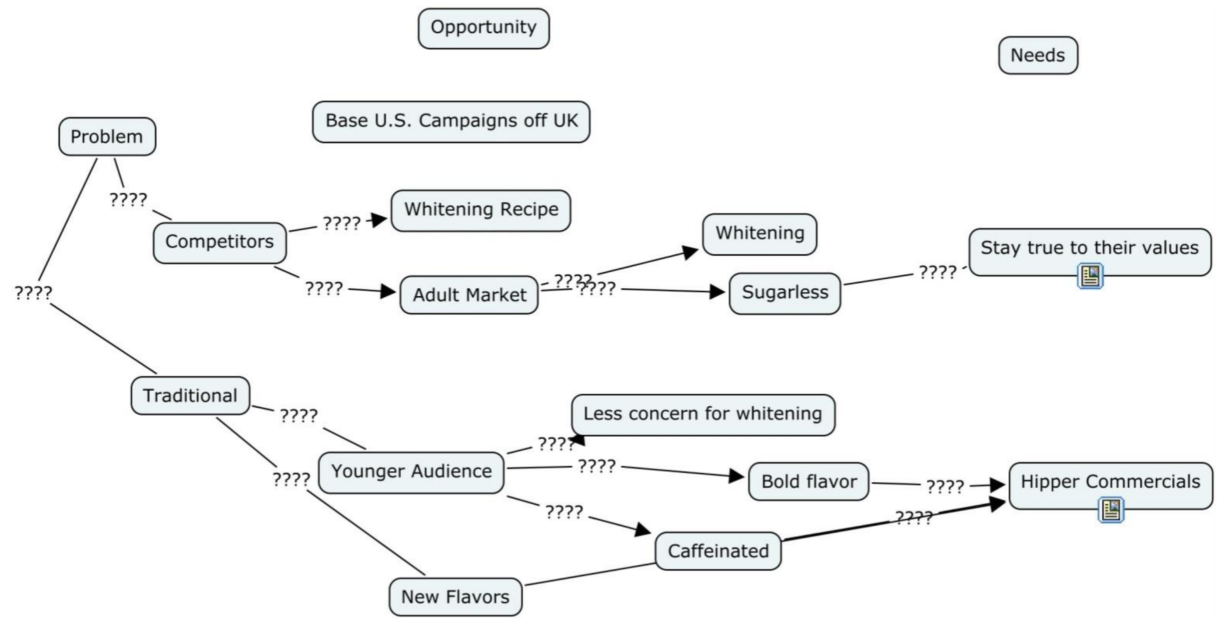
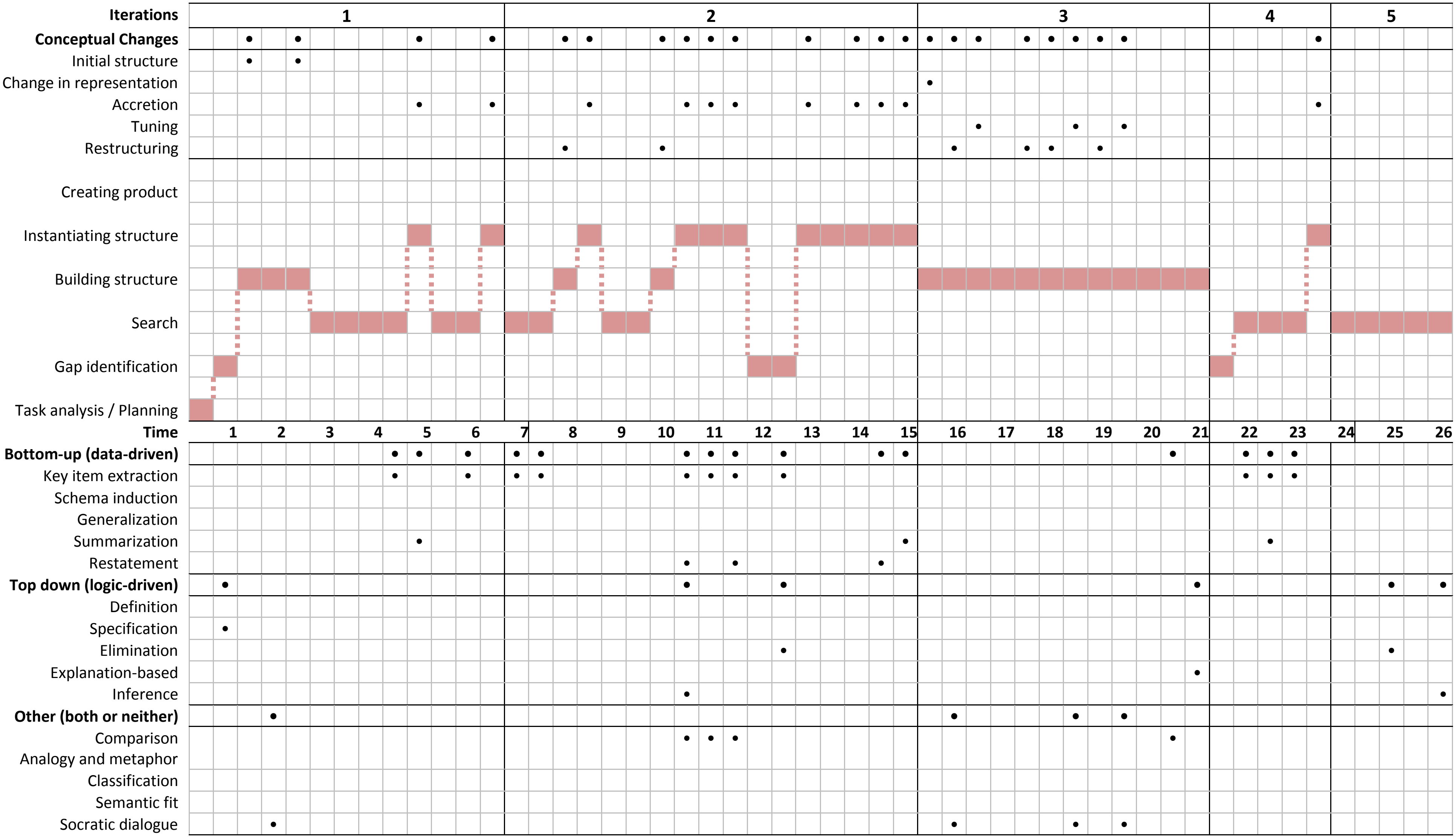
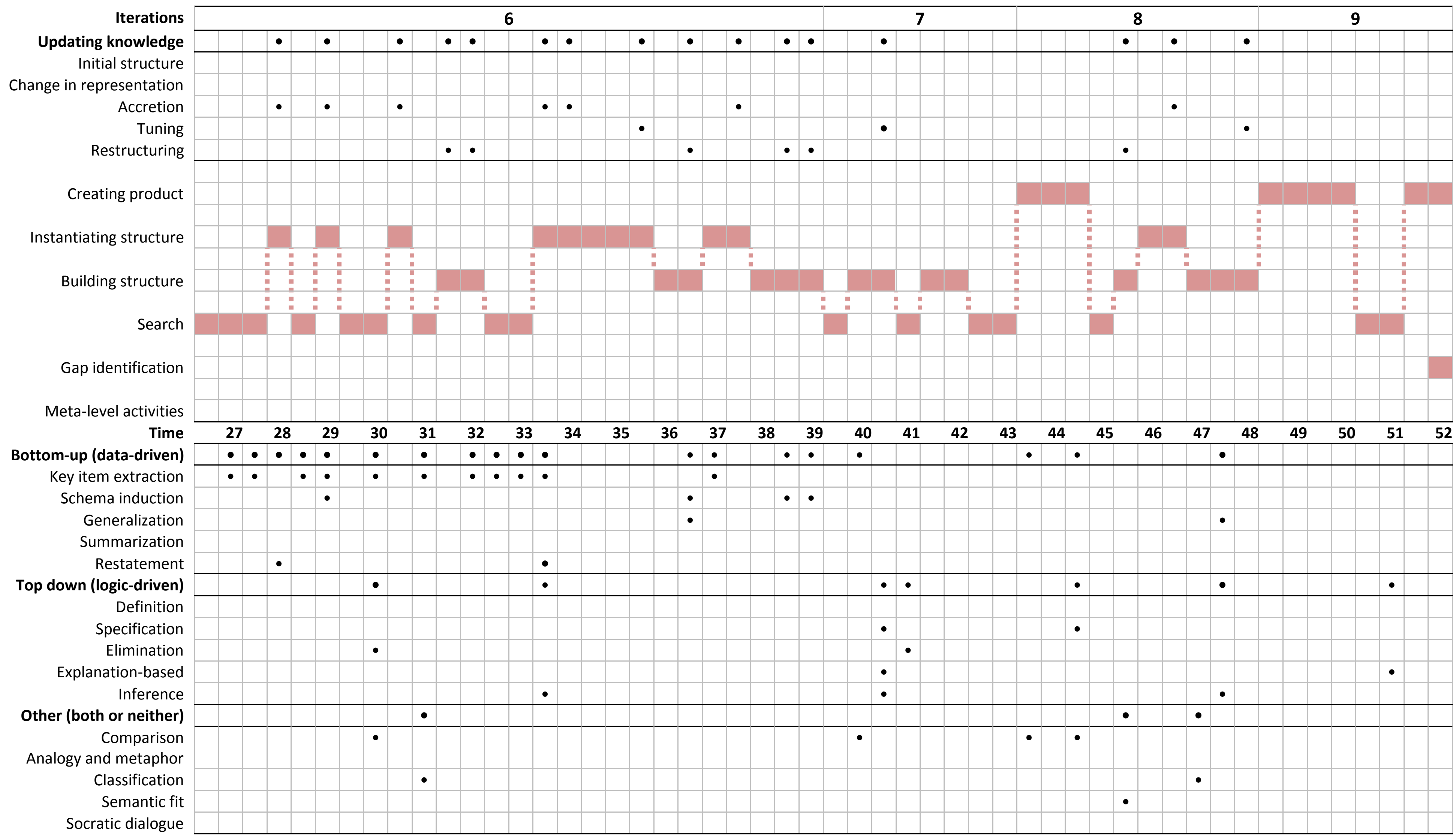
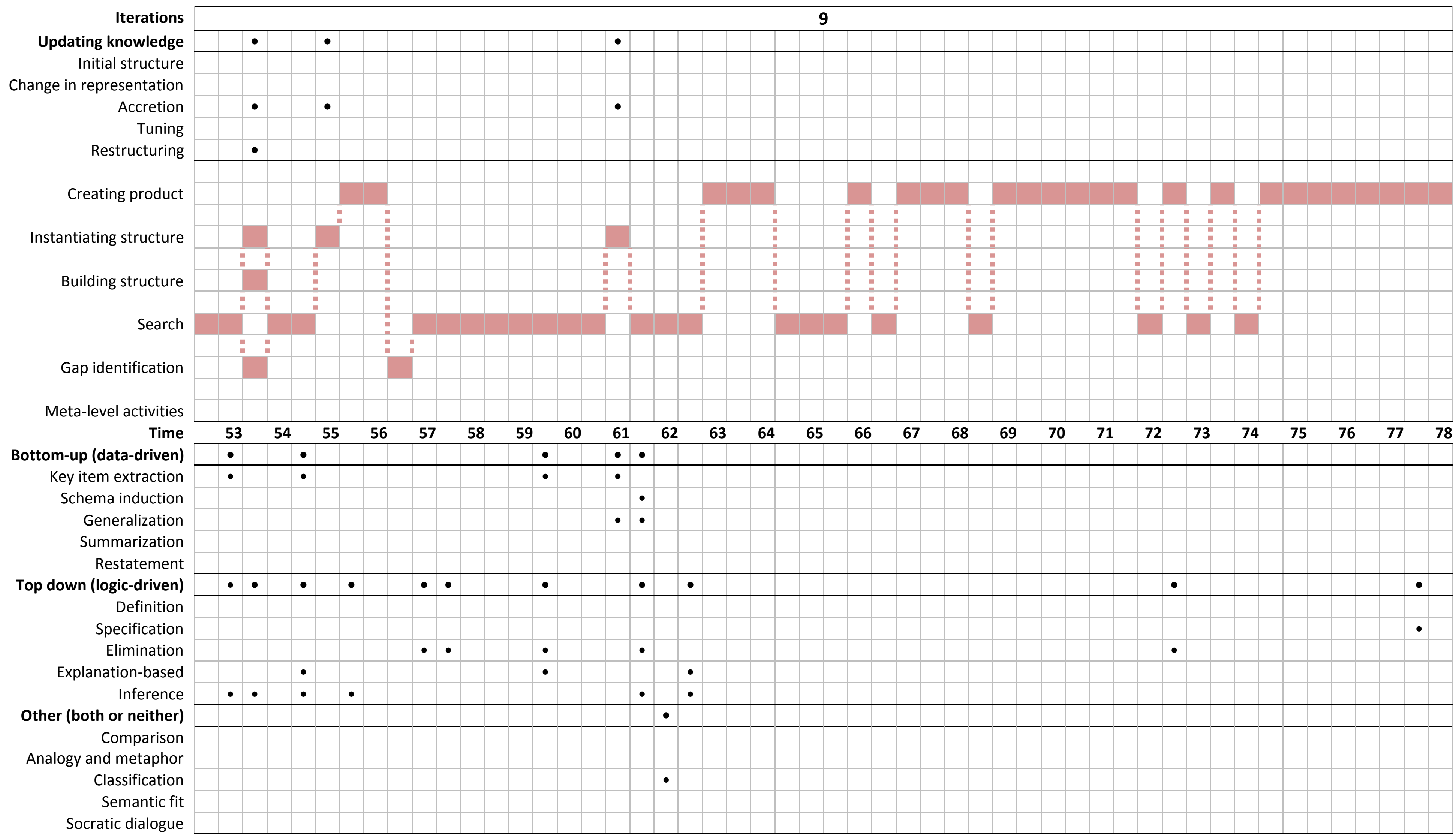


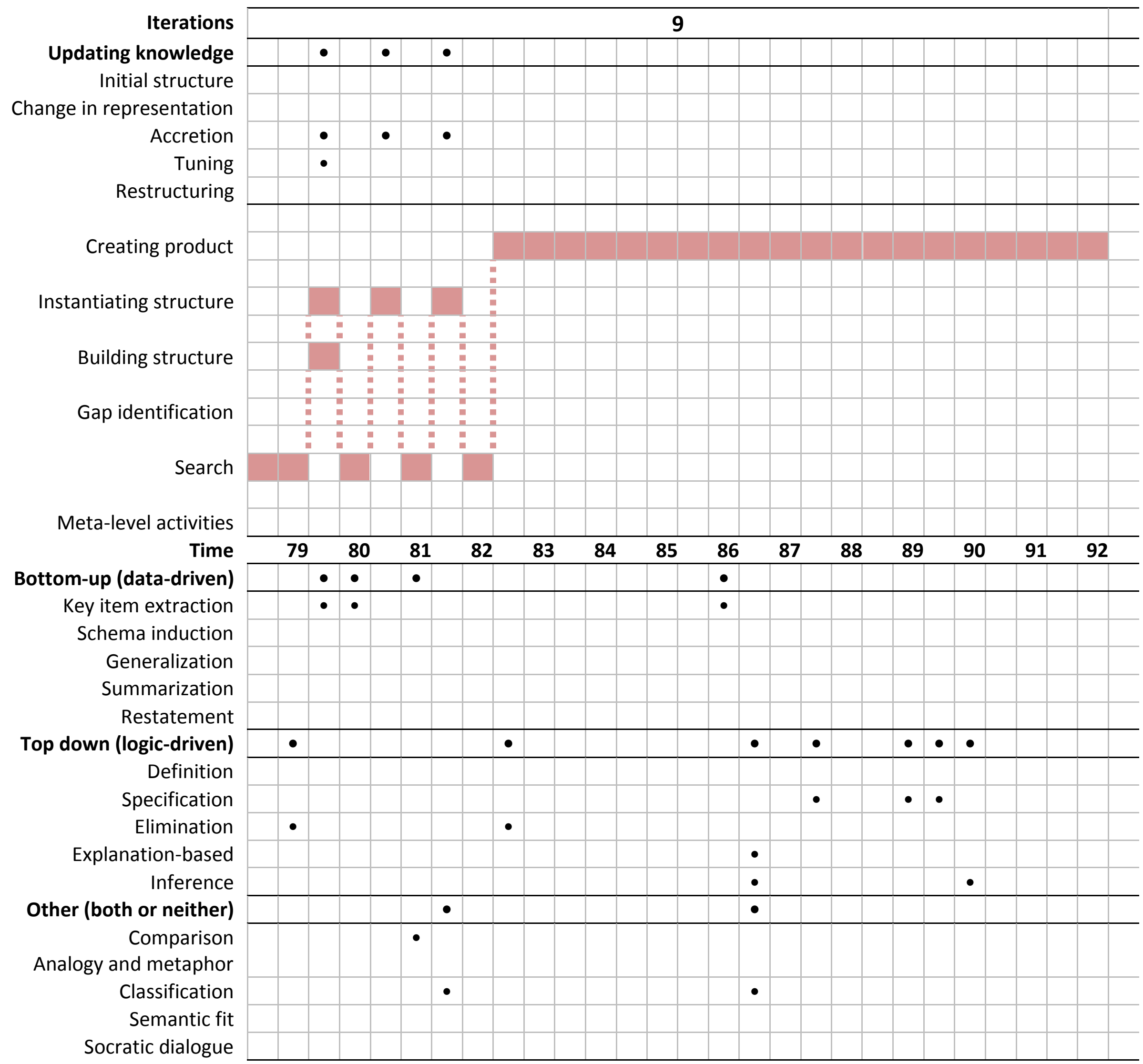
Figure E-1: Concept Map, User MB5

E.5 Sequence Chart, User MB5




Iterations
6
7
8
9
Updating knowledge
Initial structure
Change in representation
Accretion
Tuning
Restructuring
Creating product
Instantiating structure
Building structure
Search
Gap identification
Meta-level activities
Time
27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52
Bottom-up (data-driven)
Key item extraction
Schema induction
Generalization
Summarization
Restatement
Top down (logic-driven)
Definition
Specification
Elimination
Explanation-based
Inference
Other (both or neither)
Comparison
Analogy and metaphor
Classification
Semantic fit
Socratic dialogue


Iterations
9
Updating knowledge
Initial structure
Change in representation
Accretion
Tuning
Restructuring
Creating product
Instantiating structure
Building structure
Search
Gap identification
Meta-level activities
Time
53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78
Bottom-up (data-driven)
Key item extraction
Schema induction
Generalization
Summarization
Restatement
Top down (logic-driven)
Definition
Specification
Elimination
Explanation-based
Inference
Other (both or neither)
Comparison
Analogy and metaphor
Classification
Semantic fit
Socratic dialogue


Iterations
9
Updating knowledge
Initial structure
Change in representation
Accretion
Tuning
Restructuring
Creating product
Instantiating structure
Building structure
Gap identification
Search
Meta-level activities
Time
79
80
81
82
83
84
85
86
87
88
89
90
91
92
Bottom-up (data-driven)
Key item extraction
Schema induction
Generalization
Summarization
Restatement
Top down (logic-driven)
Definition
Specification
Elimination
Explanation-based
Inference
Other (both or neither)
Comparison
Analogy and metaphor
Classification
Semantic fit
Socratic dialogue

Appendix F: Energy and Election News Case, User MJ3

This appendix presents a detailed case in the study. It includes four parts:

F.1: a description of the case abridged from the assignment package provided by the user.

F.2: an example think-aloud excerpt with coding and explanations of the coding.

F.3: a case description that was derived from the coding illustrated in E.2

F.4: note pages and concept map created by the user.

F.1 Assignment Description

Energy and Election News Story
Do your own research using search engines that locate polling results as well as scholarly and journalistic sources, and write a 400-word story about the role of energy, including surrounding factors such as global warming, as an issue in the election. The story can be an overview of the issue, or you can focus the topic to a specific facet of the issue.

Table F-1: Assignment Description (Energy and Election News)

F.2 Example Think-aloud Protocol with Coding, MJ3

Time	B. User activity	C. Think-aloud Protocol	Processes	Conceptual Changes	Cognitive Mechanisms
…	B56 …	C56 …	..	…	…
47:10:: 47:34	B57 did a Google search with "PEW" to locate the PEW homepage	C57 Oh my Goodness. I am also going to look at the PEW, which is good. Gallup is the one not as good I think.	Focused search for data		
47:35:: 48:35	B58 Browsed the PEW website, did a search within the site for "energy", found an article "Overview: As Gas Prices Pinch, Support for Energy Exploration Rises"	C58 So PEW [47:30::47:41] that's the one. I am wondering if they might have like a data archive perhaps… March 08, "political survey"… probably not what I need… so, survey reports post-debate… find "energy" Aha, "energy exploration" good good good. Wow this is good. Yes! Definitely good. It is July 1st. I need to read this.	Focused search for data		Key item extraction Extracted items: "political survey", "energy exploration"
48:36:: 50:39	B59 Read the article, copied and pasted a few paragraphs from the article into the *Polling* page of the notes	C59 [SP] Um, let's see, "drilling", "partisan gap over energy exploration disappears" [48:48::49:01 reading] oh wow this is good good good. So polling… I feel good about this. [49:14::49:21] gosh this is great why did not I think of this before. [SP] This is good because it kind of talked about trends a little bit: "partisan gap over energy exploration disappears"… yes this is definitely going in [49:46::49:54] so I am just going to do the… I should put this… okay I am just going to put these because I don't want to put the graphics into the notes. [50:04::50:39]	Instantiating structure Fact "partisan gap over energy exploration disappears" linked to concept "Polling"	Accretion Fact "partisan gap over energy exploration disappears" added to users' knowledge	Key item extraction Extracted items: "drilling", "partisan gap over energy exploration disappears". Generalization generalized the fact as a trend.
50:40:: 51:26	B60 browsed the Zogby website.	C60 We should to look at Zogby too. "Trend over time". That might be good. … "Job performance"… Oh my goodness… I am going to look at archive. Don't see energy staring at me really. So I am just going to just go from here because I think I just need to really put these together…	Focused search for data, failed		Key item extraction Extracted items: "trend over time" and "job performance", not relevant.
51:27	B61 Read the first page in OneNote (untitled)	C61 now I am going to do more of the reading. [51:33::51:45] I am going to [51:50::51:55] I feel that this CMap is good… but I don't know if I really want to use it. [52:06::52:22] hmm, I need this detail about John McCain and I need this "Lexington" thing, because I think it is important.			Key item extraction Extracted item: "Lexington"

Time	B. User activity	C. Think-aloud Protocol	Processes	Conceptual Changes	Cognitive Mechanisms
52:32	B62 Titled the page "general info on energy and election 2008" which was untitled.	C62 and I need to get this a name "general info on energy and elections 2008"	Building structure Recognized concept "general info"	Tuning Added a concept	
52:51	B63 Continued to read the notes on this OneNote page.	C63 Now I am going to move to this. What is this? This is Time. This is like today. Okay, I think this is really important…	search		
53:12	B64 Created a new node in CMap "issue of the economy defeating the environment?"	C64 I am going to put it down here because I think I am going to address this… "economy"… "issue of the economy defeating the environment"….	Building structure Recognized concept "issue of the economy defeating the environment"	Tuning Added a concept	Key item extraction Extracted item: issue of the economy defeating the environment
53:38	B65 Changed the labels of the "Barack Obama" and "John McCain" nodes, adding "General History" below each, connected the Obama and McCain nodes to the three NY Times issues	C65 Okay, so yeah, I am going to do a general history (of Barack Obama and McCain), which might go down to the New York Time thing that I found because I think that will work nicely together. So this is how they stand on each one. [54:28::55:21] I don't know how I feel about this particular tool (CMap)… yes, I think this might be a good start for now.	Building structure Recognized relationships between the candidates and the issues to be discussed	Tuning Changed labels of two concepts, added a few links	Semantic fit Examined how the concepts and relationships fit with each other
55:56	B66 highlighted a few sentences from the article, read intelligibly highlighted more	C66 Okay, I want to use this idea "with the tanking economy dominating the news, and the government willing to virtually bankrupt itself to bail out the financial sector, it could be hard to push the climate change agenda – and possibly hard to find any money left to support it."… Oh! and this is kind of like global warming too. … Oh this is good because it brings in global warming. [56:49::57:00] this is a great article… I think I do not really understand completely about offshore drilling but I feel like…	Instantiating structure Fact linked to concept "issue of the economy defeating the environment" Gap identification	Accretion Added a fact to knowledge structure	Key item extraction Extracted item: issue of the economy defeating the environment
…	B67 …	C67 …	…	…	…

Table F-2: Example Think-aloud Protocol with Coding, User MJ3

F.3 Case Description of Search-Sensemaking Iterations, MJ3

Iterations	Brief Description	Paths		Conceptual changes	Cognitive mechanisms
		Search	Sensemaking		
1	Searched for general information on energy and election, found data on general info, put into her notes	Exploratory search	Instantiating structure	Accretion	Key item extraction
2	Searched for candidates' general stands on energy/environment, failed to make sense out of the article found.	Exploratory search	Attempted to build structure, but failed		Key item extraction
3	Searched for candidates' general stands on energy, adapted the articles energy factors/ issues, created a page for each candidate and put notes their positions on each issue under relevant pages. Then she identified gaps (global warming, and polling) and created a note page for each.	Exploratory search for structure	Building structure Instantiating structure Gap identification	Accretion Tuning	Key item extraction Comparison
4	Searched for global warming, but found only general information, put that into her notes	Focused search for data	Instantiating structure	Accretion	Classification
5	Searched for polling data, did not find useful polling data on energy	Focused search for data			Key item extraction
6	Browsed through her notes, decided to do an overview story instead talking about global warming, created a map outlining the story concepts, decided to talk about only three of the issues adapted from Iteration 3. Found a new lead (actual polling sites) for search.		Building structure Updating knowledge	Tuning Re-structuring	Key item extraction Specification Elimination Explanation-based mechanism Semantic fit
7	Searched two polling sites, and found actual polling data, put it into notes; noticed the issue of economy defeating energy, added that to the structure (map), added the debate into the map	Focused search for data	Instantiating structure Building structure Updating knowledge	Accretion Tuning	Key item extraction Generalization Schema induction Comparison Semantic fit
8	Wrote the story, looked for details when needed, decided not to talk about the debate after all.	Focused search for data	Instantiating structure Building structure Updating knowledge	Accretion Tuning	Key item extraction Comparison Elimination Semantic fit

Table F-3: Search-Sensemaking Iterations, User MJ3

F.4 Note Pages and Concept Map, MJ3

- General Info on Energy and Election 2008
- Barack Obama
  - Overview
  - Federal Gas Tax Holiday
  - Taxing Oil Company Windfall Profits
  - Domestic Drilling
  - Ethanol Subsidies
  - Expanding Nuclear Power
  - Coal Plants and Coal-to-Liquid Fuel
  - Statement from Obama campaign website
- John McCain
  - Overview
  - Federal Gas Tax Holiday
  - Taxing Oil Company Windfall Profits
  - Domestic Drilling
  - Ethanol Subsidies
  - Expanding Nuclear Power
  - Coal Plants and Coal-to-Liquid Fuel
  - Statement from McCain campaign website
- Polling

Table F-4: Note Pages and Structure, User MJ3

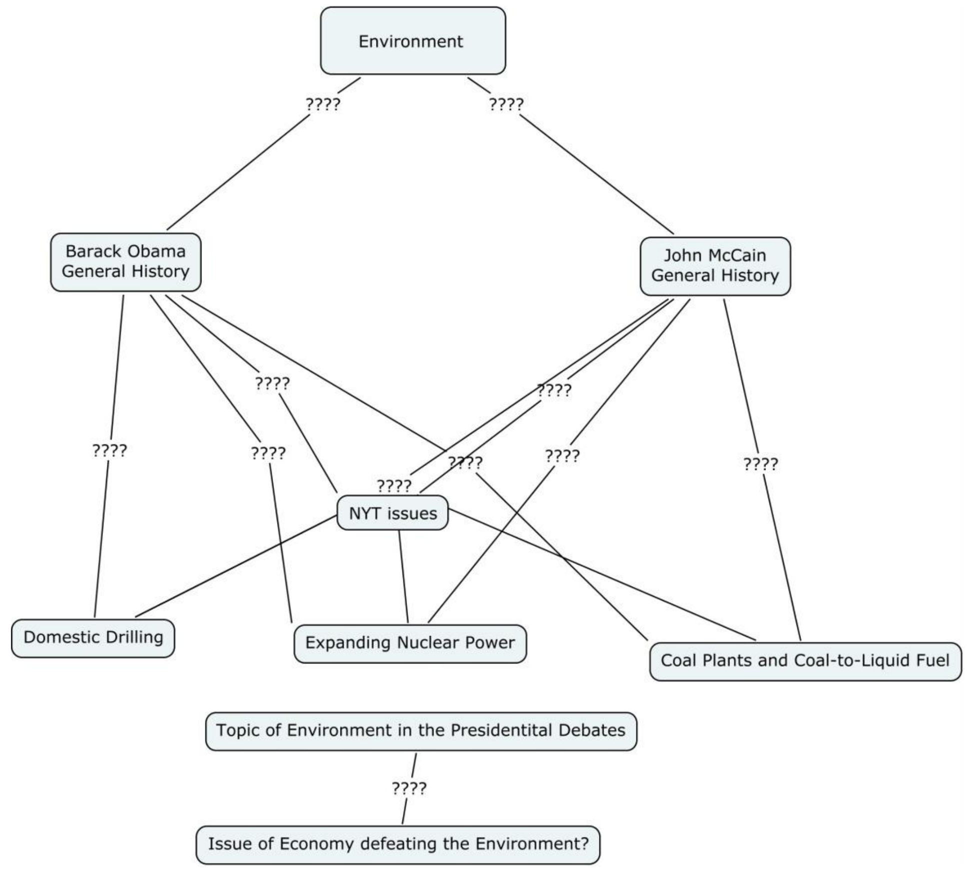


Figure F-1: Concept Map, User MJ3

Bibliography

Abbott, V., Black, J. B., & Smith, E. E. (1985). The representation of scripts in memory. Journal of memory and language, 24(2), 179-199.

Alavi, m., & Leidner, D. E. (2001). Review: Knowedge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. MIS Quarterly, 25(1), 107-136.

American Library Association (1989). Presidential Committee on Information Literacy. Final Report. from http://www.ala.org/ala/mgrps/divs/acrl/publications/whitepapers/presidential.cfm

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated Learning and Education. Educational researcher, 25(4), 5.

Anderson, R. C. (1984). Some Reflections on the Acquisition of Knowledge. Educational Researcher, 13(9), 5-10.

Arthur, W. B. (1994). Inductive Reasoning and Bounded Rationality (The El Farol Problem). The American economic review, 84(2).

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). Educational psychology. New York: Holt, Rinehart and Winston.

Baldonado, M. Q. W., & Winograd, T. (1997). SenseMaker: an information-exploration interface supporting the contextual evolution of a user's interests. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

Baron, J. (1994). Thinking and Deciding (2nd ed.): Cambridge University Press.

Basili, R., Croce, D., Cao, D. D., & Giannone, C. (2009). Learning Semantic Roles for Ontology Patterns. Paper presented at the IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. Online Information Review, 13(5), 407-424.

Bell, P. (1997). Using argument representations to make thinking visible for individuals and groups. Paper presented at the Computer Supported Collaborative Learning.

Bell, P. (2000). Scientific arguments as learning artifacts: designing for learning from the web with KIE. International Journal of Science Education, 22(8), 797-817.

Bruce, C. (1997). The seven faces of information literacy. Adelaide: Auslib Press.

Bystrom, K. (2002). Information and Information Sources in Tasks of Varying Complexity. Journal of the American Society for Information Science and Technology, 53(7), 581.

Cacioppo, J. T., & Petty, R. E. (1981). Social Psychological Procedures for Cognitive Response Assessment: The Thought-Listing Technique. In T. V. Merluzzi, C. R. Glass & H. Genest (Eds.), Cognitive Assessment (pp. 309-342). New York: Guilford Press.

Califf, M. E., & Mooney, R. J. (1999). Relational learning of pattern-match rules for information extraction. Paper presented at the Proceedings of the Sixteenth National Conf. on Artificial Intelligence.

Card, S. (2009). Information Visualization. In A. Sears & J. A. Jacko (Eds.), Human-Computer Interaction: Design Issues, Solutions, and Applications: CRC Press.

Cardie, C. (1997). Empirical methods in information extraction. AI Magazine, 18(4), 65-79.

Carley, K., & Palmquist, M. (1992). Extracting, Representing, and Analyzing Mental Models. Social Forces, 70(3), 601-636.

Chi, E. H., & Card, S. K. (1999). Sensemaking of Evolving Web Sites Using Visualization Spreadsheets. Paper presented at the Proceedings of the 1999 IEEE Symposium on Information Visualization.

Chi, M. T. H. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), Cognitive Models of Science: Minnesota Studies in the Philosophy of Science (pp. 129-186). Minneapolis, MN: University of Minnesota Press.

Chi, M. T. H. (2006). Laboratory Methods to assess the representations of experts' and novices' Knowledge. In K. A. Ericsson, N. Charness, P. Feltovich & R. R. Hoffman (Eds.), Cambridge Handbook of Expertise and Expert Performance (pp. 167-184): Cambridge University Press.

Chi, M. T. H. (2007). Three types of conceptual change:Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), Handbook of research on conceptual change. Hillsdale, NJ: Erlbaum.

Chitrapura, K. P., Joshi, S., & Krishnapuram, R. (2006). Search result summarization and disambiguation via contextual dimensions. Paper presented at the Proceedings of the 15th ACM international conference on Information and knowledge management.

Cho, K.-L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving Educational Technology Research and Development, 50(3), 1042-1629.

Choo, C. W. (2006). The Knowing Organization. Oxford, UK: Oxford University Press.

Cimiano, P., Hotho, A., & Staab, S. (2005). Learning Concept Hierarchies from Text Corpora using Formal Concept Analysis. Journal of Artificial Intelligence Research, 24, 305-339.

Ciravegna, F. (2001). Adaptive information extraction from text by rule induction and generalisation. Paper presented at the Seventeenth International Joint Conf. on Artificial Intelligence.

Cohen, W. W. (2000). Automatically Extracting Features for Concept Learning from the Web. Paper presented at the Seventeenth International Conference on Machine Learning.

Coppola, B., Gangemi, A., Gliozzo, A., Picca, D., & Presutti, V. (2009). Frame Detection over the Semantic Web The Semantic Web: Research and Applications (Vol. 5554, pp. 126-142): Springer Berlin / Heidelberg.

Creswell, J. W. (2003). Research design: Qualitative, quantitative, and mixed method approaches (Vol. 2nd). Thousand Oaks, CA: Sage Publication, Inc.

Dansereau, D. F. (1985). Learning Strategy Research. In J. Segal, S. F. Chipman & R. Glaser (Eds.), Thinking and learning skills (Vol. 1, pp. 209-239). Hillsdale, NJ: L. Erlbaum.

Das, J. P. (1994). Assessment of cognitive processes: The PASS theory of intelligence.

Denzin, N. K., & Lincoln, Y. S. (2003). Collecting and Interpreting Qualitative Materials (2nd ed.): Sage Publications, Inc.

Dervin, B. (1980). Communication gaps and inequities: Moving toward a reconceptualization. In B. Dervin & M. Voigt (Eds.), (Vol. 2, pp. 73-112). Norwood, NJ: Ablex.

Dervin, B. (1992). From the Mind's Eye of the User: the Sense-Making Qualitative-Quantitative Methodology. In J. D. Glazier & R. R. Powell (Eds.), Qualitative Research in Information Management (pp. 64-81).

Dervin, B. (1998). Sense-Making theory and practice: An overview of user interests in knowledge seeking and use. Journal of Knowledge Management, 2(2), 36-46.

Díaz, A. (2007). User-model based personalized summarization. Information processing & management, 43(6), 1715.

Doddington, G., Mitchell, A., Przybocki, M., Ramshaw, L., Strassel, S., & Weischedel, R. (2004). The Automatic Content Extraction (ACE) Program: Tasks, Data, and Evaluation. Paper presented at the LREC 2004.

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizaing Change in the Cognitive Construction of Knowledge. Educational Psychologist, 33(2/3), 109-128.

Eisenberg, M. B. (2008). Information Literacy: Essential Skills for the Information Age. Journal of Library and Information Technology, 28(2), 39-47.

Eiter, T., Ianni, G., Lukasiewicz, T., Schindlauer, R., & Tompits, H. (2008). Combining answer set programming with description logics for the Semantic Web. Artificial Intelligence, 172(12-13), 1495-1539.

Ericsson, K. A., & Simon, H. A. (1993). Protocol Analysis: Verbal reports as data. Cambridge, MA: Bradford books/MIT Press.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. Mind, Culture, & Activity, 5(3), 178-186.

Fikes, R. (1985). The role of frame-based representation in reasoning. Communications of the ACM, 28(9), 904.

Fiol, C. M., & Huff, A. S. (1992). Maps for managers: Where are we? Where do we go from here. The Journal of management studies, 29(3), 267.

Florian, R., Hassan, H., Ittycheriah, A., Jing, H., Kambhatla, N., Luo, X., et al. (2004). A Statistical Model for Multilingual Entity Detection and Tracking. Paper presented at the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics.

Florian, R., Ittycheriah, A., Jing, H., & Zhang, T. (2003). Named entity recognition through classifier combination. Paper presented at the CoNLL-2003, Edmonton, Canada.

Florian, R., Jing, H., Kambhatla, N., & Zitouni, I. (2006). Factorizing Complex Models: A Case Study in Mention Detection. Paper presented at the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition and Communications, 32(4), 365-387.

Franz, M., McCarley, J. S., Ward, T., & Zhu, W. (2001). Unsupervised and Supervised Clustering for Topic Tracking.

Gaines, B. R., & Shaw, M. L. G. (1995). WebMap: Concept Mapping on the Web. Paper presented at the WWW4.

Gersh, J., Lewis, B., Montemayor, J., Piatko, C., & Turner, R. (2006). Supporting insight-based information exploration in intelligence analysis. Communications of ACM, 49(4), 63-68.

Gick, M. L., & Holyoak, K. J. (1983). Schema Induction and Analogical Transfer. Cognitive Psychology(15), 1-38.

Grabowski, B. L. (1996). Genertive learning: past, present, future. In D. H. Jonassen (Ed.), Handbook of Research for Educational Communications and Technology (pp. 897-918). New York: Simon & Schuster Macmillan.

Graesser, A. C., & Clark, L. F. (1985). Structures and procedures of implicit knowledge. Norwood, NJ: Ablex Publishing Corp.

Gredler, M. E. (2004). Learning and Instruction: Theory into Practice (5th ed.): Prentice Hall.

Grishman, R. (1997). Information extraction: Techniques and challenges Lecture Notes In Computer Science: International Summer School on Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology (Vol. 1299, pp. 10-27). London, UK: Springer-Verlag.

Hamill, K. A., & Zamora, A. (2007). The use of titles for automatic document classification. Journal of the American Society for Information Science, 31(6), 396-402.

Hansen, P., & Bystrom, K. (2005). Conceptual framework for tasks in information studies. Journal of the American Society for Information Science and Technology, 56(10), 1050.

Hoffman, R. R. (Ed.). (1992). The psychology of expertise: cognitive research and empirical AI: Springer-Verlag New York, Inc.

Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. (1995). Eliciting knowledge from experts a methodological analysis. Organizational Behavior and Human Decision Processes, 62(2), 129-158.

Hsieh, H., & Shipman, F. M. (2002). Manipulating structured information in a visual workspace. Paper presented at the 15th annual ACM symposium on User interface software and technology.

Huang, X., & Soergel, D. (2006). An evidence perspective on topical relevance types and its implications for exploratory and task-based retrieval. Information Research, 12(1), 281.

Hyerle, D. (1996). Visual Tools for Constructing Knowledge. Alexandria, VA: Association for Supervision and Curriculum Development.

Inference (2009). Encyclopædia Britannica. Retrieved September 04, 2009, from Encyclopædia Britannica Online: http://www.britannica.com/EBchecked/topic/287533/inference

Ingwersen, P. (1992). Information Retrieval Interaction. London, UK: Taylor Graham Publishing.

Ingwersen, P., & Järvelin, K. (2005). The Turn: Integration of Information Seeking and Retrieval in Context: Springer.

Johnson-Laird, P. N. (1999). Deductive reasoning. Annual Review of Psychology, 50(1), 109.

Jonassen, D. H. (2005). Tools for Representing Problems and the Knowledge Required to Solve Them Knowledge and Information Visualization (Vol. 3426, pp. 82-94): Springer Berlin / Heidelberg.

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). Structural knowledge: techniques for representing, conveying, and acquiring structural knowledge. Hillsdale, NJ: L. Erlbaum.

Jonassen, D. H., & Henning, P. (1996). Mental models: knowledge in the head and knowledge in the world. Paper presented at the Proceedings of the 1996 international conference on Learning sciences.

Jonassen, D. H., & Wang, S. (1993). Acquiring Structural Knowledge from Semantically Structured Hypertext. Journal of Computer-Based Instruction, 20(1), 1-8.

Kambhatla, N. (2006). Minority Vote: At-Least-N Voting Improves Recall for Extracting Relations. Paper presented at the ACL '06.

Kavale, K. A. (1980). The reasoning abilities of normal and learning disabled readers on measures of reading comprehension. Learning Disability Quarterly, 3, 34-45.

Kim, S., & Soergel, D. (2005). Selecting and measuring task characteristics as independent variables. Paper presented at the 68th Annual Meeting of the American Society for Information Science and Technology.

Kirk, R. E. (1995). Experimental Design: Procedures for the Behavioral Sciences (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Klavans, J. L., & Muresan, S. (2000). DEFINDER: Rule-based Methods for the Extraction of Medical Terminology and their Associated Definitions from On-line Text. Paper presented at the AMIA.

Klein, G., Moon, B., & Hoffman, R. R. (2006a). Making Sense of Sensemaking 1: Alternative Perspectives. IEEE Intelligent Systems, 21(4), 70-73.

Klein, G., Moon, B., & Hoffman, R. R. (2006b). Making Sense of Sensemaking 2: A Macrocognitive Model. IEEE Intelligent Systems, 21(5), 88-92.

Krizan, L. (1999). Intelligence essentials for everyone. Washington, DC: Joint Military Intelligence College.

Kuhlthau, C. C. (1993). Seeking Meaning. Norwood, NJ: Ablex.

Kuhlthau, C. C. (2004). Seeking Meaning: A Process Approach to Library and Information Services (2nd ed.). Westport, CT: Libraries Unlimited.

Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: sense-making in a complex and complicated world. IBM System Journal, 42(3), 462-483.

Lee, C.-S., Kao, Y.-F., Kuo, Y.-H., & Wang, M.-H. (2007). Automated ontology construction for unstructured text documents. Data & Knowledge Engineering, 60(3), 547-566.

Lehmann, F. (1992). Semantic Networks in Artificial Intelligence. New York, NY: Elsevier Science Inc.

Lohman, D. F. (2005). Reasoning abilities. In R. J. Sternberg & J. Pretz (Eds.), Cognition and Intelligence: identifying the mechanisms of the mind (pp. 225): Cambridge University Press.

Lönneker, B. (2003). Semantic Frame Elements in Annotated Concept Frames. Paper presented at the Preceedings of the 17th International Congress of Linguists (CIL17).

Lowrance, J. D., Harrison, I. W., & Rodriguez, A. C. (2001). Capturing Analytic Thought. Paper presented at the Proceedings of the 1st International Conference on Knowledge Capture.

Luhn, L. P. (1958). The automatic creation of literature abstracts. IBM Journal of Research and Development, 159-165.

Marchionini, G. (2006). Exploratory search: from finding to understanding. Commun. ACM, 49(4), 41-46.

Markham, K. M. (1994). The concept map as a research and evaluation tool: Further evidence of validity. Journal of research in science teaching, 31(1), 91.

Maxwell, J. A. (2005). Qualitative research design: An interactive approach (2nd ed.). Thousand Oaks, CA: Sage Publication, Inc.

McKeown, K. R., Klavans, J. L., Hatzivassiloglou, V., Barzilay, R., & Eskin, E. (1999). Towards multidocument summarization by reformulation: progress and prospects. Paper presented at the sixteenth national conference on Artificial intelligence and the eleventh Innovative applications of artificial intelligence conference innovative applications of artificial intelligence (AAAI '99).

Medin, D. L., & Smith, E. E. (1984). Concepts and Concept Formation. Annual review of psychology, 35(1), 113-138.

Mikheev, A., Moens, M., & Grover, C. (1999). Named Entity Recognition without Gazetteers. Paper presented at the EACL '99.

Miles, M. B., & Huberman, M. (1994). Qualitative Data Analysis: An Expanded Sourcebook (2nd ed.): Sage Publication, Inc.

Minsky, M. (1975). A Framework for Representing Knowledge. In P. H. Winston (Ed.), The Psychology of Computer Vision (pp. 211-277): McGraw Hill.

Neches, R., Fikes, R., Finin, T., Gruber, T., patil, R., Senator, T., et al. (1991). Enabling Technology for Knowledge Sharing. AI Magazine, 12(3), 37-56.

Nenkova, A., Passonneau, R., & McKeown, K. (2007). The Pyramid Method: Incorporating human content selection variation in summarization evaluation. ACM Trans. Speech Lang. Process., 4(2), 4.

Nielsen, J., Clemmensen, T., & Yssing, C. (2002). Getting access to what goes on in people's heads?: reflections on the think-aloud technique. Paper presented at the Proceedings of the second Nordic conference on Human-computer interaction.

Nisbett, R. E., & Ross, L. (1980). Human Inference: Strategies and Shortcomings of Social Judgement: Prentice Hall.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological review, 84(3), 231.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. Organizational Science, 5(1), 14-37.

Novak, J. D. (1998). Learning, Creating, and Using Knowledge: Concept Maps As Facilitative Tools in Schools and Corporations. Mahwah, NJ: L. Erlbaum Associates.

Novak, J. D. (2006). The origins of the concept mapping tool and the continuing evolution of the tool. Information Visualization, 5(3), 175.

Noy, N. F., Fergerson, R. W., & Musen, M. A. (2000). The Knowledge Model of Protégé-2000: Combining Interoperability and Flexibility Knowledge Engineering and Knowledge Management. Methods, Models, and Tools: Proceedings of the 12th International Conference, EKAW 2000, Juan-les-Pins, France, October 2-6, 2000. (Vol. 1937, pp. 69-82): Springer Berlin / Heidelberg.

Noy, N. F., & Musen, M. A. (2000). PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment. Paper presented at the Seventeenth National Conference on Artificial Intelligence (AAAI-2000).

Pang, B., & Lee, L. (2004). A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. Paper presented at the Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics.

Perkins, D. N. (1985). General cognitive skills: Why not? In S. F. Chipman, J. Segal & R. Glaser (Eds.), Thinking and Learning Skills (Vol. 2, pp. 339). Hillsdale, NJ: Lawrence Erlbaum Associates.

Piaget, J. (1978). The Development of Thought: Equilibration of Cognitive Structures. New York: Viking Penguin.

Pirolli, P., & Card, S. K. (2005). The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis. Paper presented at the International Conference on Intelligence Analysis. from https://analysis.mitre.org/proceedings/Final_Papers_Files/206_Camera_Ready_Paper.pdf

Plous, S. (1993). The psychology of judgment and decision making. New York: McGraw-Hill.

Potter, W. J. (1996). An Analysis of Thinking and Research about Qualitative Methods. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Qu, Y., & Furnas, G. W. (2007). Model-Driven Formative Evaluation of Exploratory Search: A Study under a Sensemaking Framework. In R. W. White, G. Marchionini & G. Muresan (Eds.), Evaluating Exploratory Search Systems: Special Issue of Information Processing and Management (Vol. To appear).

Rajan, K., Ramalingam, V., Ganesan, M., Palanivel, S., & Palaniappan, B. (2009). Automatic classification of Tamil documents using vector space model and artificial neural network. Expert Systems with Applications, 36(8), 10914-10918.

Ram, S. (2004). Semantic conflict resolution ontology(SCROL): an ontology for detecting and resolving data and schema-level semantic conflicts. IEEE transactions on knowledge and data engineering, 16(2), 189.

Ramage, D., Heymann, P., Manning, C. D., & Garcia-Molina, H. (2009). Clustering the tagged web. Paper presented at the Proceedings of the Second ACM International Conference on Web Search and Data Mining.

Riloff, E., & Schmelzenbach, M. (1998, Aug, 1998). An Empirical Approach to Conceptual Case Frame Acquisition. Paper presented at the Sixth Workshop on Very Large Corpora, Montreal.

Romano, N. C. J., Bauer, C., Chen, H., & Nunamaker, J. F., Jr. (2000). The MindMine comment analysis tool for collaborative attitude solicitation, analysis, sense-making and visualization. Paper presented at the 33rd Annual Hawaii International Conference on System Sciences.

Rumelhart, D. E., & Norman, D. A. (1981a). Accretion, tuning, and restructuring: Three modes of learning. In J. W. Cotton & R. Klatzky (Eds.), Semantic factors in cognition (pp. 37-90). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rumelhart, D. E., & Norman, D. A. (1981b). Analogical processes in learning. In J. R. Anderson (Ed.), Cognitive skills and their acquisition (pp. 335-359). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, NJ: Lawrence Erlbaum Associates.

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). The cost structure of sensemaking. Paper presented at the Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems.

Savolainen, R. (2006). Information use as gap-bridging: The viewpoint of sense-making methodology. Journal of the American Society for Information Science and Technology, 57(8), 1116.

Sharma, N. (2007). Sensemaking handoff: theory and recommendations. Paper presented at the CHI '07 extended abstracts on Human factors in computing systems.

Sharp, H., Rogers, Y., & Preece, J. (2007). Interaction Design: Beyond Human Computer Interaction. Hoboken, NJ: Wiley.

Sokal, R. R. (1974). Classification: Purposes, Principles, Progress, Prospects. Science, 185, 1115-1123.

Sowa, J. F. (Ed.) (2006) Encyclopedia of Artificial Intelligence. Wiley.

Sparck Jones, K. (1999). Automatic Summarizing: Factors and Directions. In I. Mani & M. Maybury (Eds.), Advances in Automatic Text Summarization (pp. 1-12). Cambridge, MA: MIT Press.

Staff, C., & Bugeja, I. (2007). Automatic classification of web pages into bookmark categories. Paper presented at the Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval.

Stake, R. E. (1995). The Art of Case Study Research. Thousand Oaks, CA: Sage Publications.

Stefik, M. J., Baldonado, M. Q. W., Bobrow, D., Card, S., Everett, J., Lavendel, G., et al. (1999). The knowledge sharing challenge: The sensemaking white paper: PARC, Inc.

Sternberg, R. J. (1986). Toward a Unified Theory of Human Reasoning. Intelligence, 10(4), 281.

Storrer, A., & Wellinghof, S. (2006). Automated detection and annotation of term definitions in German text corpora. Paper presented at the LREC 2006.

Suthers, D. D. (1998). Representations for Scaffolding Collaborative Inquiry on Ill-Structured Problems. Paper presented at the Annual conference of the American Educational Research.

Suthers, D. D. (1999). Effects of Alternate Representations of Evidential Relations on Collaborative Learning Discourse. Paper presented at the 1999 conference on Computer support for collaborative learning.

Suthers, D. D. (2001). Towards a systematic study of representational guidance for collaborative learning discourse. Journal of Universal Computer Science, 7(3), 254-277.

Suthers, D. D. (2003). Representational Guidance for Collaborative Inquiry. In J. Andriessen, M. Baker & D. D. Suthers (Eds.), Confronting Cognitions: Arguing to Learn (pp. 1-17): Kluwer Academic Publishers.

Tjong Kim Sang, E. F., & De Meulder, F. (2003). Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition. Paper presented at the CoNLL '03.

Toulmin, S., Rieke, R., & Janik, A. (1979). An introduction to reasoning. New York: Macmillan Publishing Co., Inc.

Vakkari, P., & Hakala, N. (2000). Changes in Relevance Criteria and Problem Stages in Task Performance. Journal of Documentation, 56(5), 540-562.

Vosniadou, S. (1989). Analogical Reasoning and Knowledge Acquisition. In S. Vosniadou & A. Ortony (Eds.), Similarity and Analogical Reasoning (pp. 413): Cambridge University Press.

Vosniadou, S., & Brewer, W. F. (1987). Theories of Knowledge Restructuring in Development. Review of educational research, 57(1), 51-67.

Vosniadou, S., & Ortony, A. (1989). Similarity and Analogical Reasoning. London, UK: Cambridge University Press.

Wang, W., & Haake, J. (1997). Supporting user-defined activity spaces. Paper presented at the Proceedings of the eighth ACM conference on Hypertext.

Wertheimer, M. (1938). Laws of Organization in Perceptual Forms. In W. Ellis (Ed.), A source book of Gestalt psychology (pp. 71-88). London: Routledge & Kegan Paul.

Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. Paper presented at the Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing.

Witte, R., & Bergler, S. (2007). Fuzzy Clustering for Topic Analysis and Summarization of Document Collections. Lecture Notes in Computer Science, 4509, 476.

Wittrock, M. C. (1990). Generative Learning Processes of Comprehension. Educational Psychologist, 24(4), 345-376.

Wright, W., Schroh, D., Proulx, P., Skaburskis, A., & Cort, B. (2006). The Sandbox for Analysis - Concepts and Methods. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems.

Yin, R. K. (2003). Case Study Research: Design and Methods (3rd ed.): Sage Publication, Inc.

Zhang, J. (1997). The Nature of External Representations in Problem Solving. Cognitive Science, 21(2), 179.

Zhang, J. (2000). External representations in complex information processing tasks. In A. Kent (Ed.), Encyclopedia of Library and Information Science. New York: Marcel Dekker, Inc.

Zhou, G., & Su, J. (2001). Named entity recognition using an HMM-based chunk tagger. Paper presented at the Proceedings of the 40th Annual Meeting on Association for Computational Linguistics.

Zhou, L. (2007). Ontology learning: state of the art and open issues. Information Technology and Management, 8(3), 241-252.