

# DESIGNING THEORY-BASED SYSTEMS: A CASE STUDY

John B. Smith and Marcy Lansman

Department of Computer Science  
University of North Carolina  
Chapel Hill, NC 27599-3175  
919-962-1792 919-962-1979

jbs@cs.unc.edu lansman@cs.unc.edu

## ABSTRACT

In this paper, we discuss principles for designing and testing computer systems intended to support users' thinking as they perform open-ended or ill-defined tasks. We argue that such systems inherently and inevitably implement a model of users' cognitive behaviors. Making that model explicit can provide system developers with guidance in making design decisions. However, both model and system must be tested and refined. We discuss these principles in relation to a case study in which our group developed a hypertext-based writing environment and then tested that system in a series of experimental studies of writers' strategies.

**KEYWORDS:** system design, cognitive modes and strategies, cognitive models, task analysis, user testing

## INTRODUCTION

*What should be the relationship between human-computer interaction studies and the design and testing of actual systems?*

Few would disagree that results of human-computer studies should play a larger role in the design of computer systems. However, we have not seen very many instances where this has actually happened. There are many reasons for this. Most HCI studies have addressed the ways individuals interact with existing systems. Many have been concerned with specific interface issues -- such as representation, layout, use of color, ease of operation for specific commands, etc. -- rather than broad, patterns of behavior that might be more useful for developers. As a result, much of this work remains unknown to system designers or it has been incorporated piecemeal.

One promising development can be seen in recent theoretical discussions that sketch broad approaches to system design, frequently drawing on work from other disciplines including speech act theory [Winograd &

Flores, 1986], activity theory [Bodker, 1989], ethnography [Suchman, 1987]. These discussions have provided useful and convincing insights into the situated activities of users; but because of their generality, applying these insights to specific design problems is often difficult. For example, some writers describing the Scandinavian approach suggest that factors as far-afield as the medieval social structure of the nation are relevant for contemporary system designers [Floyd, *et. al.*, 1989].

Thus, much of the work in HCI is either too specific or too general to provide practical guidance for system building. What designers need is *practical* guidance that addresses *overall* system design. Such a method would enable them to reliably and predictably build "whole" systems that have internal consistency and integrity.

During the past seven years, our research group has been engaged in a program of research that relates to these issues. We did not begin with the intention of addressing broad issues of design methodology; rather, we were interested in users' cognitive strategies for a particular task -- technical and scientific writing -- and in building a hypertext-based computer system to support that task. We soon realized, however, that to understand this task would require basic research in cognitive theory and, to be useful, the theory would have to relate in specific ways to the support system we were building.

We have also come to realize that the *method* applies to many tasks other than writing. In our own project, we have extended this approach to the tasks of software development and to collaborative work. But, we believe the method applies to any complex, open-ended, or ill-structured task. Included in this category would be design, planning, extended problem-solving and other tasks that require sustained human thinking that is carried out with the help of a computer system.

Such systems have sometimes been called *intelligence augmenting* or *intelligence amplifying* software [Engelbart, 1973]. An example is hypertext systems in which users represent abstract structures of ideas as networks of nodes and links -- each node corresponding to a concept and each link corresponding to an association or relationship. The computer is said to augment or amplify *intrinsic* human conceptual processes by representing a larger set of ideas than that which can be held in short

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

term memory and by permitting more numerous and more complex traversal operations than would otherwise be possible. IA systems can be contrasted, on the one hand, with systems that enable users to manipulate data structures in accord with *extrinsic*, rather than *intrinsic*, rules -- for example, accounting systems. They can be contrasted, on the other hand, with *artificial intelligence* systems that, by design, seek to simulate human thought, but are intended to operate autonomously rather than cooperatively.

While intelligence amplifying systems were once primarily a curiosity, we believe they will become the predominant form of software for microcomputers and workstations. Indeed, we already see a trend in this direction in multiple window multitasking operating systems, in systems that integrate multiple tools, in the proliferation of hypertext systems, and in the more powerful tools that are appearing for a variety of design tasks.

Building systems that can help users think more efficiently and more effectively requires new methods for their design as well as new methods to test and refine those designs. It seems self-evident to us that a system that seeks to help its users carry out complex, open-ended tasks will be more successful if it implements a *model* of that process than if it does not. Such a model will include a data, or product, component and a set of system operations to affect that product, but it must emphasize the *cognitive* processes and strategies that are used by human beings to build or to understand the complex conceptual structure that is being represented in that product. For example, a good deal is known about the cognitive processes that writers use to produce technical, scientific, and other kinds of expository documents. However, the task is too complex and subject to too many undetermined variables for someone to produce a set of rules that would automatically generate a well-written document. Nevertheless, we could describe at a more abstract level a set of *steps* that can help a human writer select and organize material, plan the document, edit its sentences, etc. And we could build a system that, by design, could help its users carry out those steps. But both the enactment of the steps and the control of the supporting computer system are dependent upon the intellect of the human user, as opposed to an independent algorithm.

This notion of a *set of steps* suggests what we mean by a *model* for open-ended tasks; in the section that follows, we explain this concept more precisely. By a *theory-based system*, we mean a system that is consistent with a particular model for a complex, open-ended task. After the discussion of theory, we describe a general approach for designing such systems, illustrated by one particular system developed by our group. Finally, we discuss issues concerned with testing and refining both system and model.

## THEORY

Systems intended to amplify or augment their users' thinking inevitably include in their design a model of those mental activities. This model is inherent in the way conceptual objects are represented, in the operations provided to manipulate them, in options for moving objects from one context to another, etc. But, regardless of whether system designers are aware of it or not, the model is there, embedded in the implementation.

A more effective approach, we believe, is for system designers to be consciously aware of this model. In an ideal world, it would be defined before system design begins; however, in practice, the model is likely to be worked out in parallel with system design and refined after users start working with the system.

Over the past seven years, our group has followed this approach in developing several systems. The models we have developed have been expressed in terms of a set of *cognitive modes* and the *strategies* and *tactics* human beings use to engage the various modes and to move among them. In this section, we first discuss these concepts in general terms; after that, we illustrate them by describing a particular set of modes used for technical and scientific writing. We wish to emphasize, however, that modes and strategies are general concepts that can be used to model a number of intellectual tasks in addition to writing.

We define a cognitive mode to be a particular way of thinking used to accomplish a particular *goal*, that will be realized by producing a particular kind of *product*, drawing on particular cognitive *processes*, in accord with a particular set of *constraints*. The *product* produced in a mode is the symbolization of a concept or relationships among concepts. Different cognitive modes provide different options for representing concepts or structures -- for example, words, diagrams, notes, outlines, and other forms. Thus, different forms prevail in different modes. Cognitive *processes* act on products to define them or to transform one form into another. Thus, certain processes are favored in certain modes, while others are de-emphasized or suppressed. The *goal* of a mode is the individual's intention for engaging that particular way of thinking. While goals are abstract, they are made concrete in the particular product the individual aims to produce in that mode. *Constraints* determine the choices available in a mode. Constraints are relaxed or tightened in accord with the individual's large-scale strategies for engaging different modes of thinking for different purposes.

To illustrate these concepts, consider two modes frequently used by expository writers: exploratory thinking and organizing. During *exploration*, the goal is to externalize ideas, consider different combinations, and to gain a general sense of the information available or missing. Thus, constraints are minimal to encourage creativity and multiple perspectives. The processes that are

emphasized are recalling from memory, associating, relating, and building small component structures. The products generated are, thus, notes, jottings, diagrams, perhaps loose networks of ideas. During *organization*, the goal is to plan the actual document to be written; thus, constraints are tightened to produce a logical, coherent organizational plan. That plan is normally expressed as a hierarchy or some other regular form. And the processes are analyzing, synthesizing, sustained conceptual building, and refinement based on noting consistent/inconsistent relations in the structure. *Exploration* and *organization* are, thus, distinctly different ways of thinking. And they differ still from other activities such as actual writing and several forms of editing.

Modes are used *strategically*. With respect to writing, experienced writers use the various modes in accord with a general strategy they know and use to accomplish a particular intellectual activity. But they also switch modes for tactical reasons that arise during the course of work. By strategy, we mean writers' overall understanding of the writing process and the steps they have learned that enable them to get their writing done. By tactics, we refer to the fact that writers shift from one mode to another in response to specific problems that occur. Researchers studying writers' strategies have noted that writers may return to exploration mode after an organizational phase when they realize during writing that the plans they produced earlier are inadequate [Hayes & Flower, 1980]. Thus, modes help writers focus their attention on one set of activities at a time, while strategies provide them with an overall sense of direction as well as the means to resolve problems that arise during the process.

When writers use cognitive modes in accord with a global strategy, they are likely to produce a series of related intermediate products. For example, during exploration some writers represent concepts externally, cluster them, and then link them into a loose network of associations. During organization, they transform that loose network of ideas into a coherent structure for the document. During writing, the individual concepts and relations in the organizational plan are transformed into continuous prose, graphic images, or other developed forms. During editing, they refine the structure and expression of the draft document. Thus, writers produce a *flow* of intermediate products in which the output of one mode becomes the input for another.

However, this flow of products is not one-way and continuous. Rather, as writers shift modes iteratively and recursively to solve problems, the flow of intermediate products goes back and forth, as well. For example, writers may find while organizing that they do not have critical information needed for a particular section. Rather than interrupt their thinking to get that information, they may elect to continue but leave the section undeveloped. Later, when the missing information is available, they may interrupt their writing, revert to organization or perhaps even exploration mode, and build the missing portion of the document's structure. When the missing

piece has been filled in, they resume writing. [Smith & Lansman, 1991]

The concept of cognitive mode can be applied to tasks other than writing. Over a range of intellectual activities, individuals divide tasks into subtasks, set goals and subgoals, produce intermediate objects or subassemblies, and employ different processes during different phases of the work. They do these things whether they are working with physical objects or with information objects. In considering this broader range of tasks, we have found it useful to differentiate between the *general* notion of mode and the *specific* modes used for a given task. The first -- defined as the interdependent configuration of goal, products, processes and constraints -- can be viewed as an *architectural construct*. While different tasks draw on different modes, they can all be described within the general framework of modes and strategies.

A *model*, as we use the term, is a particular set of modes and the rules that account for the relationships among them. Thus, a model of writing identifies a particular set of modes -- those used by particular writers under particular writing conditions -- and the particular strategies and tactics used by those writers, which are learned and/or developed by those writers. Consequently, we should not be surprised to see different groups of writers using different sets of modes for a given task. Some sets of modes and strategies, it can be argued, are preferable to others because they are more efficient or suit certain groups better than other sets.

We developed a model of writing that emphasizes separate and abstract planning modes; we call this approach the *strategic method for writing* [Smith & Smith, 1987]. The seven modes included in the strategic method, along with their constituents, are summarized in Figure 1. Since a key step in the methodology we are discussing is mapping model to system design, we describe the different modes of this model in more detail here. In the section that follows, we show how these particular cognitive modes are mapped to various system modes.

*Exploration* mode is used to gain a general sense of the material available for the document. During exploration the writer retrieves ideas for the text from memory and from source materials, writes them down as short phrases, clusters them and notes specific relations among them. Some people write phrases on "post-its" and cluster the post-its to represent the relationships between ideas. Other writers may express their ideas as rough sentences and link these sentences with arrows. The products of exploration mode are always intermediate, i.e., these products do not show up directly as part of a draft. There are few constraints on the form of the products of exploration. For example, ideas need not be expressed in complete sentences. There is, however, at least one constraint on the process: writers avoid evaluating the ideas that come up during exploration and, instead, simply record them.

*Situational Analysis* mode is used to identify aspects of the rhetorical situation that affect the text. As was the case for exploration, there are few constraints on the

	<b>Processes</b>	<b>Products</b>	<b>Goals</b>	<b>Constraints</b>
<b>Exploration</b>	<ul style="list-style-type: none"> <li>•Recalling</li> <li>•Representing</li> <li>•Clustering</li> <li>•Associating</li> <li>•Noting subordinate-superordinate relations</li> </ul>	<ul style="list-style-type: none"> <li>•Individual concepts</li> <li>•Clusters of concepts</li> <li>•Networks of related concepts</li> </ul>	<ul style="list-style-type: none"> <li>•To externalize ideas</li> <li>•To cluster related ideas</li> <li>•To gain general sense of available concepts</li> <li>•To consider various possible relations</li> </ul>	<ul style="list-style-type: none"> <li>•Flexible</li> <li>•Informal</li> <li>•Free expression</li> </ul>
<b>Situational Analysis</b>	<ul style="list-style-type: none"> <li>•Analyzing objectives</li> <li>•Selecting</li> <li>•Prioritizing</li> <li>•Analyzing audiences</li> </ul>	<ul style="list-style-type: none"> <li>•High-level summary statement</li> <li>•Prioritized list of readers (types)</li> <li>•List of (major) actions desired</li> </ul>	<ul style="list-style-type: none"> <li>•To clarify rhetorical intentions</li> <li>•To identify and rank potential readers</li> <li>•To identify major actions</li> <li>•To consolidate realization</li> <li>•To set high-level strategy for document</li> </ul>	<ul style="list-style-type: none"> <li>•Flexible</li> <li>•Extrinsic perspective</li> </ul>
<b>Organization</b>	<ul style="list-style-type: none"> <li>•Analyzing</li> <li>•Synthesizing</li> <li>•Building abstract structure</li> <li>•Refining structure</li> </ul>	<ul style="list-style-type: none"> <li>•Hierarchy of concepts</li> <li>•Crafted labels</li> </ul>	<ul style="list-style-type: none"> <li>•To transform network of concepts into coherent hierarchy</li> </ul>	<ul style="list-style-type: none"> <li>•Rigorous</li> <li>•Consistent</li> <li>•Hierarchical</li> <li>•Not sustained prose</li> </ul>
<b>Writing</b>	<ul style="list-style-type: none"> <li>•Linguistic encoding</li> </ul>	<ul style="list-style-type: none"> <li>•Coherent prose</li> </ul>	<ul style="list-style-type: none"> <li>•To transform abstract representation of concepts and relations into prose</li> </ul>	<ul style="list-style-type: none"> <li>•Sustained expression</li> <li>•Not (necessarily) refined</li> </ul>
<b>Editing: Global Organization</b>	<ul style="list-style-type: none"> <li>•Noting large-scale relations</li> <li>•Noting and correcting inconsistencies</li> <li>•Manipulating large-scale structural components</li> </ul>	<ul style="list-style-type: none"> <li>•Refined text structure</li> <li>•Consistent structural cues</li> </ul>	<ul style="list-style-type: none"> <li>•To verify and revise large-scale organizational components</li> </ul>	<ul style="list-style-type: none"> <li>•Focus on large-scale features and components</li> </ul>
<b>Editing: Coherence Relations</b>	<ul style="list-style-type: none"> <li>•Noting coherence relations between sentences and paragraphs</li> <li>•Restructuring to make relations coherent</li> </ul>	<ul style="list-style-type: none"> <li>•Refined paragraphs and sentences</li> <li>•Coherent logical relations between sentences and paragraphs</li> </ul>	<ul style="list-style-type: none"> <li>•To verify and revise coherence relations within intermediate sized components</li> </ul>	<ul style="list-style-type: none"> <li>•Focus on structural relations among sentences and paragraphs</li> <li>•Rigorous logical and structural thinking</li> </ul>
<b>Editing: Expression</b>	<ul style="list-style-type: none"> <li>•Reading</li> <li>•Linguistic analysis</li> <li>•Linguistic transformation</li> <li>•Linguistic encoding</li> </ul>	<ul style="list-style-type: none"> <li>•Refined prose</li> </ul>	<ul style="list-style-type: none"> <li>•To verify and revise text of document</li> </ul>	<ul style="list-style-type: none"> <li>•Focus on expression</li> <li>•Close attention to linguistic detail</li> </ul>

Figure 1: Seven cognitive modes for writing.

products of situational analysis. These products include notes, lists, and diagrams that represent what is known or assumed about potential readers. These products are intermediate and are used to guide decisions made in other modes. During situational analysis, the writer envisions potential readers, establishes priorities among them, imagines what readers know about the subject matter and decides how he would like the text to affect them. Thus, situational analysis allows the writer to make consideration of context a conscious conceptual process.

*Organization* mode is used to develop a single, coherent structure for the text. The writer uses the ideas and component structures produced in exploration mode as the raw materials for organization mode. He groups sets of ideas under their logical superordinate headings, generating those superordinate concepts when necessary. He breaks other ideas down into their components. He experiments with various organizational schemes to determine which one fits the rhetorical goals he has developed during situational analysis. The processes required by organization involve examining the logical relationships between ideas. The product, for the writer, is a hierarchical structure containing three to four levels of topic headings. The organizational process is constrained by the requirement that the result be a single organizational scheme which includes all the major ideas that are to appear in the text.

*Writing* mode is used to translate the ideas in the organizational scheme into sentences. The product is a rough draft. Both the organizational scheme and the rules of English grammar constrain that product. While writers vary widely as to the quality of the prose they expect to produce in writing mode, the writer strives for a first draft that is grammatical and rhetorically suitable for the purposes established during situational analysis. But he anticipates making major structural and linguistic changes during the editing phases.

Editing is done in three different cognitive modes. During *global editing*, the writer addresses the large-scale structure of the document. The purpose is to make sure that the document as a whole makes the right point, that the right parts are present, and that they are in the right order. The primary constraint is that attention is focused on the high-level, structural features of the document and that the details be ignored. During global editing, the writer evaluates large-scale relations, notes logical inconsistencies among the parts of the document, and corrects or manipulates these large, structural components. The product is a refined version of the document - one that has a sharper central focus than the original draft and one in which the large components fit together more comfortably.

*Coherence editing* requires the writer to shift attention to intermediate-sized units of the text, such as paragraphs and sections. The purpose is to examine the logical, sequential order of, first, the paragraphs within sections and, then, sentences within paragraphs. The primary constraint is again to focus attention on units of a particular size. Cognitive processes include evaluating

coherence relations and restructuring paragraphs to make relations clear. Some sentences may have to be transformed or rewritten to make them fit together. The product is a document in which the sentences and paragraphs have clear, logical relations to one another and advance the larger purpose of the section they comprise.

*Expression editing* represents still a different mode of thinking. Whereas coherence editing is concerned with sentences as discrete objects to be verified and arranged, expression editing is concerned with the insides of sentences - their clarity, directness, and appropriateness for the rhetorical purposes of the document. Thus, expression editing requires close attention to linguistic detail. The processes emphasized are reading, linguistic analysis, linguistic transformation, and linguistic encoding. The product is a more refined document - one with crafted prose.

The seven cognitive modes described here represent one particular model of the writing process. Our subsequent studies suggest to us that it accurately accounts for the behavior of many writers, but not all. Thus, as is the case for any phenomena, alternative models are possible and, indeed, may be required if we are to account for the behavior of most individuals for open-ended tasks, such as writing. In a later section, we return to this issue of alternative models, but, first, we show how this particular model served as the basis for designing a writing support system.

## SYSTEM DESIGN

We have suggested above that writing can be viewed as a complex process involving different cognitive modes, such as the particular set of modes just described. A key question for system design, then, is how best to support these different cognitive modes and the flow of intermediate products among them? We will try to answer this question here, both in general but also in more detail for the Writing Environment (WE) developed by our group.

Two basic designs are possible. In a single mode system, all system functions would always be available. For writing, the set of functions would be the union of those required to support all of the cognitive processes for the different cognitive modes discussed above. By contrast, a multimodal approach would divide the environment into separate system modes, each corresponding to one or more of the cognitive modes. If the second approach were followed, each system mode would include only the functions appropriate for its corresponding cognitive mode(s).

We adopted a multimodal system design for WE for several reasons. As we discussed in the previous section, writers manage the overall writing task by dividing that process into phases in which they engage different cognitive modes. Each mode is unique in terms of its particular combination of processes, products, goals, and constraints. Consequently, supporting these large-grained

"chunks" of activity, each with its own unique requirements, in separate system modes seemed both natural and efficient: natural, in that system architecture would both mirror and reinforce cognitive strategy; efficient, in that specific system operations could be matched closely with specific cognitive processes and with specific intermediate products developed by writers during the task.

We made this design decision recognizing that it ran counter to widely-held beliefs that systems with multiple user interface modes are less desirable than so-called *seamless* systems that profess to have only a single, all-inclusive interface mode. The case against multimodal interfaces is strongest for systems that are controlled through textual commands. Problems arise when input from the keyboard is interpreted as textual data in one mode or context and as commands to the system in other contexts. The problem is compounded when the system includes multiple control modes, resulting in different command interpretations for the same input. However, the problem with multiple mode systems has always been making the user aware of what mode the system is in at any given moment. A simple remedy is available for graphics-based systems, such as the Macintosh and the other multiple window systems, in the form of visual cues that signal the mode. In fact, one can reasonably argue that today's multiple window multitasking operating systems are inherently multimodal; they are not perceived to be such -- and, indeed, have even been called *seamless* -- because they have solved the multimodal problem so effectively that users are unaware they are switching modes when they switch between applications/windows.

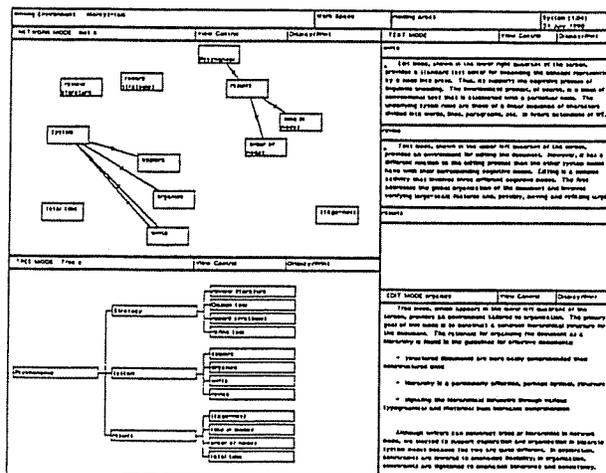


Figure 2: Writing Environment (WE). Overview of the four system modes: Network, Tree, Editor, and Text Modes

Building theory-based systems is relatively straightforward when the task model is expressed in terms of *modes*. For

example, the Writing Environment presents the user with four system modes that correspond with six of the seven cognitive modes included in the *strategic* model, outlined in Figure 1. The default layout of the screen is shown in Figure 2; however, the individual modes can be resized freely. The upper left window, called *network* mode, is intended for *exploration*. The underlying data model in this mode is a directed graph embedded in a two-dimensional space. Thus, the user has maximum flexibility with which to represent concepts as nodes (boxes with a word or phrase to express the idea), move them to form clusters of loosely related ideas, and link them to denote more specific relationships. Small conceptual structures can also be built here and used later in the other modes.

To support the *organization* cognitive mode in which the user builds the actual plan for the document, the system provides a *tree* mode, shown in the lower left corner of Figure 2, in which the user is constrained to create a hierarchical structure. (Reading comprehension research has shown that a hierarchical document is likely to be more easily and more accurately comprehended than documents with other structures [e.g., Kintsch & van Dijk, 1978; Schwartz & Flammer, 1981; Ausubel, 1963; Meyer, Brandt, & Bluth, 1980; Kieras, 1980].) While users could have continued working in network mode to build a tree, they are encouraged to shift to the system mode specifically intended for organization. Thus, system design encourages, but does not demand, users to transform their (loose) network of ideas into a well-formed hierarchical structure.

At any time in the process they can open a node and write or edit a block of text that will be associated with that node. *Editor* mode, shown in the lower right corner, is a conventional text editor. There, writers transform into text the concepts and the relations among them represented in the graph or tree. Eventually, we will provide other editors so that they may express an idea as a drawing. In fact, the general framework of the system is sufficiently general so that it would permit sound, video, or other forms of expression so long as editor and display functions are available.

Finally, the upper right system mode, called *text* mode, is intended for *coherence editing*. While the tree represents the structure of the document and the logical sequence of nodes or blocks of text that comprise it, text mode constructs a linear form of the implied text by stepping through the tree -- top to bottom, left to right. In it, one can see transitions from the text in one node to the text in another node, move sentences from one to another, etc. Eventually, we will replace text mode with a full WYSIWYG editor, but none is currently available that we can use.

In summary, the four system modes correspond with *exploration*, *organization*, *writing*, and *coherence editing*. For *structure editing*, writers uses tree mode: by moving branches and nodes around in the tree, users reorganize the

text of the associated document. To support *expression editing*, writers may use either editor or text modes. Thus, six of the seven cognitive modes shown in Figure 1 are supported by the four system modes in WE. At present, WE does not support *situational analysis* mode. We have developed several heuristics that help writers with this thinking process [Smith & Smith, 1987], but since the products produced in this mode of thinking do not become literal parts of the document, we have not built a system mode to support it.

## STUDIES

The third step in theory-based design is to test and refine both system and theory. To refine the model for writing and the WE system that was based upon it, we carried out a series of experimental studies under quasi-naturalistic conditions. Those studies, in addition to serving this purpose, also address a broader set of cognitive and human-computer interaction issues concerned with writers' cognitive strategies and patterns of behavior. Conducted over a 3-4 year period, they produced two different kinds of information that bear on the validity of the *strategic* model and the WE system design. First, we collected comments from participants in the form of responses on written questionnaires completed after several days of system use as well as oral responses made during debriefings. Second, as we explain in more detail below, we collected objective data of users' actions with the system in the form of machine-recorded protocols. By examining these data, we can see quite clearly and concretely which task or system operations caused users problems. Features that cause difficulty -- which we label *turbulence* to indicate interference with the natural flow of information and intent between user and system-- suggest inconsistencies either between the model and users' actual cognitive processes and strategies or between the model and its realization in the system design.

In the remainder of this section, we discuss these data in more detail.

The gist of our approach is this: a user works with an application system we have developed or modified to produce a machine-readable transcript of all *actions* -- rather than *keystrokes* -- performed by that user during the session. That session can take place in the laboratory or in the user's natural working environment. The data can be used to recreate, or *replay*, an approximation of the original session, but in a fraction of the original time. It can also be analyzed automatically by one of the *cognitive grammars* we have developed. These grammars constitute models of users' cognitive strategies for a given task using a particular computer system. The grammars are used to parse the protocols, producing a parse tree that is a concrete representation of a particular user's strategy for a given session. While these parse trees can be examined directly, more often they are further processed by *filter programs* that count various symbols or combinations of symbols in accord with a particular analytic perspective; these derived data are then passed to a statistical utility for

conventional analysis. Finally, these various data are presented to the researcher through a combination of *static* and *animated display tools* to facilitate visualization and interpretation. This methodology has been discussed in more detail in [Smith, Smith, & Kupstas, 1991].

While experiments differed in their particular designs, all took the same basic form. The overall purpose of the experiments was to have different populations of writers use the WE system to plan and write one or more documents approximately two-three pages in length, based upon reading materials supplied to them and addressing particular rhetorical situations that identified purpose, readers, etc. Subjects came to the lab for a series of two to five half-days, normally in the same week. They spent one or two half-days learning the system by completing a structured tutorial and several sample tasks. On the third and subsequent days, subjects planned, wrote, and edited documents using the WE system. Each writing task took two-three hours, during which the system automatically (and unobtrusively) recorded detailed action-level protocols, as discussed above. This design, thus, produced the two kinds of data noted above -- comments and observable patterns in the machine-recorded protocols.

### Responses to Questionnaires

At the end of each study, we asked participants to fill out a long questionnaire about their reactions to the Writing Environment. For example, we asked them what they liked and didn't like about the system, how it compared to other word processing systems they had used, and how useful they found each of the system modes. Responses indicated that participants adapted easily to the multimodal nature of the system and to the fact that different system modes appear in different windows on the screen. They particularly liked the fact that they could see the organizational structures of their papers in a separate window as they wrote text. Here are some users' comments:

*The multiple window display is the most useful feature. The ability to see the organization of the document while editing a node is unique.*

*I like having the outline section right next to the text section for quick reference.*

We studied questionnaire responses for evidence that the system modes of the Writing Environment matched or did not match the cognitive modes of users. Responses suggested that the match was good for Network and Tree Modes, but not good for Edit and Text Modes.

Many users reported that they used Network Mode as we intended - for jotting down ideas and investigating the relationships between them. In response to the question, "Was Network Mode useful?" they wrote:

*Yes because it allowed you to get your ideas down without having to organize them.*

*Yes - I liked the 'linking' idea.*

*Yes - it was useful to get many different ideas down quickly w/o having to worry about order.*

*Yes. Different ideas could be scattered in the beginning and then connected later.*

They also found Tree Mode a useful tool for organizing their articles. In response to the question, "Was Tree Mode useful?" they said:

*Yes because one could organize your ideas from the network modes and decide which ideas were useable and which were not.*

*Yes. It was very easy to move whole connected areas of thought.*

An occasional user felt the Network and Tree Mode were redundant, but a greater number reported that they used the two modes for different types of thinking.

Responses to Edit and Text Mode were less enthusiastic. Some users liked the fact that Edit Mode encouraged them to focus on one idea at a time:

*Perfect! It kept you focused on one aspect of your paper and helped you move on in writing. I think this is what cut the time consuming task of writing.*

But many of the comments on Edit and Text Mode suggested that users had a hard time coordinating the planning modes and the writing modes to produce a text that flowed easily from section to section.

Some users wanted to use the structure they had produced in Tree Mode as a guide rather than as a fixed framework for their papers.

*I do not like being constrained in the tree portion to having each topic in the text portion. I would like to be able to write a whole outline and pick and choose which topics will be paragraphs in the text.*

*I didn't like the fact that every part of the tree diagram acquired it's own title heading. I therefor couldn't include items to a topic which didn't include substantial text also.*

Others wanted to see more of the text at one time while editing:

*It would maybe be nice to have a screen where differentiation between nodes could be suppressed. - See flow of paper and transitions.*

Some of the responses to Edit and Text Mode may reflect the fact that the Writing Environment did not have the polished editing capabilities of modern commercial word processors. But some of them also indicate that there is a mismatch between the way writers edit their texts and the way Edit and Text Modes of the Writing Environment are designed.

Users' comments clearly indicate that they want something like a WYSIWYG editor, in which they can see the text as a single unit. The challenge will be to create a new system mode which, on the one hand, allows users to see and modify the text as a whole and, on the

other hand, allows users to rearrange sections of text by moving nodes in a tree.

### Machine-Recorded Protocols

While the questionnaires reflect users' subjective reactions to the four system modes of the Writing Environment, the computer-generated protocols give us an objective view of how writers used these system modes. They tell us, for example, how users distributed their time among the various modes. Computer protocols indicated that almost all users spent a significant amount of time in both Network and Tree Modes and did work in both modes. The same is not true of Edit and Text Modes. A number of writers used either Edit or Text Mode exclusively for both writing and revising. Like the questionnaires, the protocol data indicate a good match between the cognitive processes of exploration and organization and the design of Network and Tree Modes but a poorer match between the cognitive processes of writing and revising and Edit and Text Modes.

Several aspects of the protocol data suggest that writers' strategies changed as they became more familiar with the system. During the early stages of practice, they spent a large proportion of their time experimenting with Network and Tree Modes. They often built very large trees, larger than necessary in some cases. As one user commented:

*I should have made a simple hierarchy in the tree mode and not made so many nodes. It got to a point where I would have less than a sentence in each node. Better to have much larger chunks in each node.*

Later in practice, users spent less time in the planning modes. In the one experiment, in which subjects learned to use the Writing Environment on Days 1 and 2 and wrote separate articles on Day 3 and Day 4, the amount of time spent in Tree Mode decreased significantly from Day 3 to Day 4 as did the number of nodes in the final tree.

Writing strategy also varied with the writers' knowledge of the topics they were writing about. In one study, we asked graduate students in art history and in chemistry to write two articles each, one on a particular type of Japanese Art and one on a type of metal alloy. Users spent more time *planning* their articles when they were writing on an unfamiliar topic than when they were writing on a familiar topic. But they spent more time writing and revising when they were working on the familiar topic.

One of the most interesting characteristics of the data was the pattern of alternations between the structural planning modes and the writing modes of the system. Many teachers advise their students to plan their papers first, by writing an outline, and then to use their outlines to guide their writing. According to these teachers, composing a document should take place in separate stages - first planning, then writing. We used our computer-generated protocols to ask whether our participants used the strategy teachers recommend.

We found that there was huge variability in the extent to which participants finished their planning in Network and Tree Modes before they began to write and revise in Edit and Text Modes. Figure 3 shows two extreme cases. For each subject, two broken horizontal lines represent time spent planning (in Network and Tree Modes) and time spent writing and revising (in Edit and Text Modes). The writer at the top did almost all his planning before he began to write. The writer at the bottom alternated between planning and writing throughout the session. The other 15 writers in this experiment were spread out along a continuum between these two extremes. In general, computer protocols show far more alternation between the planning the writing/revising modes of the system than we had anticipated. This finding emphasizes the need for smooth transitions between system modes.

These observations and data do not conclusively "prove" the theory or its realization in the system design. But

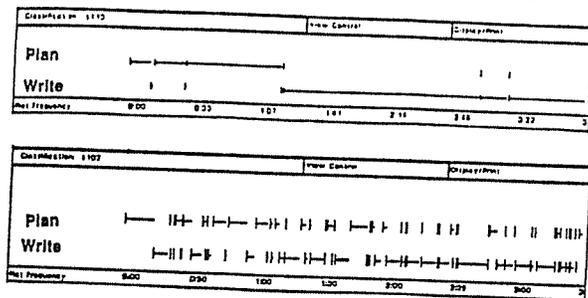


Figure 3: Each panel of this figure shows how an individual subject distributed his time between planning (Network and Tree Modes) and writing/revising (Edit and Text Modes). Time since the beginning of the session is shown on the horizontal axis. Each vertical tick represents the beginning of a planning or a writing/revising episode. The length of the horizontal line attached to the tick represents the duration of the episode. The top panel represents a writer who did almost all of his planning before he began writing and revising. The panel at the bottom represents a writer who alternated often between planning and writing/revising.

they do provide evidence that supports some design decisions while indicating the need to alter others. For example, the system's planning modes seem to closely match writers cognitive planning modes. On the other hand, the evidence suggests that although coherence editing may be a valid cognitive mode, the system mode intended to support that kind of thinking, which we called text mode, was unsatisfactory.

### CONCLUSION AND FURTHER RESEARCH

The theory-based method described here is both comprehensive with respect to system design while also providing developers with guidance in making decisions regarding specific functions and their organization within the interface. It also provides ways to test both system

and theory that can produce specific suggestions for changes that could make the system better fit its users' thinking. We have used this method to build and test a hypertext system for expository writing, and we are extending it to two other tasks/systems -- for software development and for collaborative work. While our work to date convinces us of both the efficacy and generality of the approach, much research still needs to be done.

Modes and strategies are really architectural components. Further work is needed to elaborate a more complete cognitive architecture based on them. Similarly, we hope to see other tasks analyzed within this architecture. But to do so, better observational and analytic tools are needed to study patterns of users' behavior under naturalistic conditions and extending over extended periods of time -- months rather than hours -- if we are to take into account longitudinal and adaptive effects. And we need better interface development tools that facilitate building multimodal systems in which user function can easily be edited and reorganized in order to test different modal configurations or to match configuration with particular groups of users.

This program of research will require multidisciplinary teams and will involve basic research in both cognitive science and computer science as well as in HCI. But it promises us computer systems that more closely match the way we think. This is likely to be an increasingly important concern as computing becomes both more universal and more distributed.

### ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (Grants # IRI-8519517 and # IRI-8817305), The Army Research Institute (Contract # MDA903-86-C-345), and The Office of Naval Research (Contract # N00014-86-K-0680). A number of individuals have contributed both ideas and effort to the work described here; they include Gordon Ferguson, Steve Weiss, Dick Hayes, Catherine Smith, Matt Barkley, Paulette Bush, Rick Hawkes, John Hilgedick, Hong Li, Mark Rooks, Doug Shackelford, Yen-Ping Shan, Oliver Steele, John Walker, and Irene Weber.

### REFERENCES

- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York: Grune & Stratton.
- Bodker, Susanne (1989). A human activity approach to user interfaces. *Human-Computer Interaction*, 4,3,171-195.
- Engelbart, D. C., Watson, R. W., & Norton, J. C. (1973). The augmented knowledge workshop. *AFIPS conference proceedings*, pp. 9-21.
- Floyd, C., Wolf-Michael, M., Reisin, F-M., Schmidt, G., & Wolf, G. (1989). Out of Scandinavia: Alternative approaches to software design and system development. *Human-Computer Interaction*, 4,4, 253-349.

- Hayes, J. R. & Flower, L. S. (1980). Identifying the organization of the writing process. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive Processes in Writing*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 3-30.
- Kieras, D. E. (1980). Initial mention as a signal to thematic content in technical passages. *Memory and Cognition*, 8, 345-353.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- Meyer, G. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: key for reading comprehension of ninth grade students. *Reading Research Quarterly*, 1, 72-103.
- Schwartz, M. N. K., & Flammer, A. (1981). Text structure and title-effects on comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 20, 61-66.
- Smith, J. B., & Lansman, M. L. (1991). *Cognitive modes and strategies for writing*. Chapel Hill, NC: UNC Department of Computer Science Technical Report # TR91-047.
- Smith, J. B., & Smith, C. F. (1987). *A strategic method for writing*. Chapel Hill, NC: UNC Department of Computer Science Technical Report # TR87-024.
- Smith, J. B., Smith, D. K., & Kupstas, E. (1991). Tools and techniques for automated protocol analysis. Chapel Hill, NC: UNC Department of Computer Science Technical Report # TR91-034.
- Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.
- Winograd, T., & Flores, C. F. (1986). *Understanding computers and cognition: A new foundation for design*. Norwood, NJ: Ablex.