



## **Using information systems while performing complex tasks: an example from architectural design**

ERICA DE VRIES\*

*Faculty of Technology Management, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands*

TON DE JONG

*Faculty of Educational Science and Technology, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands*

*(Received 25 March 1994, revised manuscript received 26 June 1996 and accepted 2 August 1996)*

Nowadays, information systems, such as hypertexts, allow a variety of ways in which to structure information. Information systems are also used for an increasing number of purposes. In our study we examined two different purposes for using information systems in the context of a real task: architectural design. In design processes, information gathering plays a different role depending on design phase, and both exploration and finding information are important sub-processes. A study is presented in which the main goal was to determine whether there are advantages of certain information structures when carrying out a particular activity. Both tasks and variables used in the experiment were related to characteristics of the design process as identified in the literature. Two different kinds of information gathering behaviour, browsing and searching were studied, and criterion tests corresponded to the desired outcome in the design context, i.e. enlargement of the users' information span for browsing and their knowledge of specific topics for searching. Results showed a number of interactions between information structure and information gathering task depending on the particular criterion test examined. Inspection of each purpose on its own criterion test revealed the merits of network structures for browsing, but also for searching. These results parallel the recommendations emanating from design disciplines to provide information structures for design that do not impose a hierarchical structure, but that show the complexity of design information. Recommendations for hypertext research include the study of navigational characteristics in combination with tasks and variables that are meaningful in a specific context.

© 1997 Academic Press Limited

### **1. Introduction**

Information systems are used to store large amounts of information that can be accessed through a human-computer interface. Both the way in which the information is stored, or structured, and the interface are considered to be crucial to the success of an information system (Marchionini & Schneiderman, 1988). There are, however, no general guidelines for structuring the information and the interface since their optimal design depends upon the task for which the information system

\* Present address: COAST Research Team, UMR-GRIC, CNRS-University of Lyon Lumière, ENS de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France. email: erica.de.vries@ens-lyon.fr

is used. There has been an emergence of new types of information systems (access aspect) and new types of uses of these systems (task aspect). The issue of adapting the information system to a particular task has therefore become an important one. This paper reviews some of the research on the relation between access aspects and task aspects. Major flaws can be detected in this type of research. We will describe these problems, and then propose our approach in which a task context as a whole is considered. Hereafter, we present an empirical study demonstrating how access and tasks aspects can be dealt with using the general context of the task as a whole, in our case: the (architectural) design process.

### 1.1. SEARCHING INFORMATION AND INFORMATION STRUCTURE

One of the most and most intensively studied tasks with information systems is searching information. Traditional information systems, e.g. databases using keyword-based search, are mostly designed for fact finding. Users of such systems want to find a particular fact, and to find it quickly. In a number of studies, comparisons are made of different systems on their performance in search tasks (see Dumais, 1988 for an overview). Performance is measured in terms of speed (task completion time or the time needed to answer a question, or find a fact) and accuracy (number of errors, or irrelevant items found). Characteristic of these studies is the target specificity: the information to be found is well defined and the finding of all other information is regarded inefficient.

On the access side, there has been a shift from traditional information retrieval systems to flexible information systems, such as hypertext. In hypertext, information is put into “nodes” that are related through “links”. The nodes in a hypertext can be linked in different ways to produce different structures. The structure of a hypertext determines the information that can be chosen at any particular moment. Hypertext systems have also been studied for searching information (see for example Edwards & Hardman, 1989; Simpson & McKnight, 1990; McKnight, Dillon & Richardson, 1990). Results from these studies already hint upon a relationship between structure of the information and the specific character of the task. In some situations, a traditional linear structure was found to be the most accurate and the most efficient in answering questions (e.g. McKnight *et al.*, 1990). Others found that a high speed of question answering is obtained with hierarchical structures, but that these structures may be inefficient when an answer involves relations between pieces of information (Edwards & Hardman, 1989; Mohageg, 1992). Network structures produced inconsistent results. Whereas Mohageg did not find a speed advantage for relational tasks (questions that, according to expert judgements, are easier with a relational network) using a network structure, Wright and Lickorish (1990) found that page navigation (which also involves jumping between pieces of information) resulted in higher speed for questions involving relations between text pages.

In most of these studies, question answering is the main task and performance is measured in terms of speed and accuracy. However, a difficulty in integrating the results of the studies is the large range of structures and performance measures employed. Moreover, the types of questions used vary considerably from study to study and may have little to do with consultation of an information system in real-life contexts.

Information systems, and especially hypertext, can be used for other tasks than searching information, such as information exploration. They are thought to facilitate both problem structuring and problem solving (Begoray, 1990). In fact, on the task side, the development of hypertext systems coincides with a shift in interest from searching and locating a specific fact to tasks involving browsing and exploring subject areas. The large number of possibilities for linking in information systems immediately evokes the question of how to structure hypertext information to ensure an optimal performance for the latter type of tasks.

## 1.2. BROWSING INFORMATION AND INFORMATION STRUCTURE

A number of studies have tried to link information structure to the purpose for using the information system. Different purposes are distinguished mainly to the degree to which there exists a definite target in the mind of the user. Such a dimension has also been called target orientation (Waterworth & Chignell, 1991), referring to the cognitive state the user is in when entering a system. An issue addressed in this type of research is, for example, the identification of the task situations under which each of a number of linking structures excels (Mohageg, 1992).

Mohageg (1992) tried to match task situations with different linking structures. The structures were linear, hierarchical-organizational, network-relational, and a combination of these latter two. *Relational links* are used to connect points or regions in the text thus forming a network. Two nodes may be linked because they contain contradictory information or because one node supports information from another. *Organizational links* are used to arrange nodes according to some rule; for example some hierarchical organization of information (see also Conklin, 1987). Tasks were classified using expert judgements according to whether the tasks were more easily solved using relational or hierarchical links, resulting in *relational* and *hierarchical* tasks. The main task, both the relational and the hierarchical, was question answering. Dependent variables were (amongst others) task completion time. The results showed that task completion time was lower in the hierarchy and combination condition as compared to the linear and network condition. This was observed both for relational and hierarchical tasks. So, although experts judged relational tasks easier to solve with network links, no advantage of the network condition was observed.

Waterworth and Chignell (1991) induced browsing by giving questions like "Choose four sentences from the text that best capture the essence of the subject" (p. 44). Browse questions were compared with questions encouraging a query strategy. An example of a query question is "What is the title of the Conklin article published in 1987?" (p. 43). Access to the information was also varied: using the system involved either navigation (following links) or searching (querying). Subjects had to answer both types of questions in both types of access. Waterworth and Chignell report on subjective ratings of difficulty and on task completion time. The results showed that browse questions were more difficult than query questions as indicated by the subjective ratings and also by longer completion times. Navigation was judged more difficult than querying. This effect was also visible in the task completion times, but only approached significance. However, as Waterworth and Chignell (1991) mention in their discussion, it seems paradoxical to encourage

browsing and then taking success as the speed with which an answer can be found. If browsing means exploring a subject area, success would be indicated by long completion times and by the interesting information found along the way.

### 1.3. METHODOLOGICAL ISSUES

The above cited studies reveal two large problems in studies that relate tasks and information access when the task is exploration (or browsing) of the information. The first problem is the choice of an appropriate experimental task that encourages subjects to browse the information in an information system. The second problem concerns the development of adequate dependent variables for measuring the success of a browse task. A similar problem has been identified in memory research where the importance of considering relations between tasks and tests was stressed by Bransford, Franks, Morris and Stein (1979). According to Bransford *et al.* the “goodness” of particular acquisition activities can be defined only in relation to particular testing contexts. They argue that a statement about the value of a particular acquisition activity must include reference to the nature of the testing context. Although Bransford *et al.* discuss the notion of *transfer-appropriate processing* in the area of memory research, their ideas are valid in the area of human–computer interaction as well. In other words, the shift in access and task aspects has to be followed by a reconsideration of performance tests. In studying tasks such as the exploration of a subject area, performance can no longer be examined using question answering tasks and speed and accuracy tests.

A starting point to find an escape from this situation is to look at the context in which the exploration task is taking place. The meaning of success largely depends on the context or the task situation. As a consequence, the task as a whole has to be examined in order to derive appropriate performance tests. In order to evaluate the fruitfulness of such an approach, we set out to study the use of an information system in a particular task situation. For this aim, we chose to study the design process, because designing is a complex activity in which a large amount of information is needed during the entire process. In the present study, we examined the suitability of different linking structures in relation to information gathering that takes place in different phases in the design process. Before presenting this empirical study, we provide the necessary background on design processes.

## 2. Information system use in context: the design process

Design as a problem solving process receives a growing attention from the side of cognitive scientists. Recently, the term generic design was introduced to refer to the study of design as a subject matter in its own right, independent of specific tasks or disciplines (Goel & Pirolli, 1992). Goel and Pirolli argue that design problems and tasks display similarities across various disciplines, whereas significant differences exist in the structure of design problems and non-design problems. Building upon Newell and Simon’s (1972) general theory of problem solving, Goel and Pirolli examine design in terms of the task environment and the problem space. One of the salient characteristics of the task environment of design tasks is what Goel and Pirolli (1992) call “distribution of information” which means that the designer lacks

information on for example the start state and the required end state of the design problem. These characteristics of the task environment have as a consequence that design problem solving is characterized by *problem structuring* before problem solving and the appearance of *distinct problem-solving phases*, starting with preliminary design, followed by a refinement and detailed design (Goel & Pirolli, 1992, p. 405).

One of the disciplines taken by Goel and Pirolli (1992) as an exemplar of a design task is architecture. Architectural design tasks have been the subject of a number of studies, both from a more applied viewpoint (Lawson, 1990; Wade, 1977) as well as from a cognitive perspective (Akin, 1986; Hamel, 1990). Problems in architectural design are generally accepted as ill-defined at all three aspects of a problem: the start state, the methods to be used, and the end state. One of the key issues in solving architectural design problems, and in design problems in general, is that design involves satisfying a large number of constraints. The constraints are often “open” in the sense that the constraint nor its desired level of satisfaction are explicitly stated in the problem description. Reitman (1965) described open constraints as attributes (of the goal state) whose definition includes one or more parameters the values of which are left unspecified as the problem is given to the problem solver. For example, when the architectural design problem is “design a house”, attributes can be the “size of a house” or the “shape of a house”. Such attributes and their values are not necessarily stated in the problem description, or as Goel and Pirolli (1992) put it, design constraints are not constitutive to the task. Constraints are important for structuring the initial problem space, and for narrowing down the number of possible solutions. They permit to divide a design problem into manageable sub-problems, and finally, constraints provide criteria for judging possible solutions. Some authors even state that designing can be seen as a process of expressing and exploring constraints and trying to achieve objectives (Gross, Ervin, Anderson & Fleisher, 1988). An important consequence is that different problem solvers in a domain can vary in the number and the kind of constraints that they consider important, and the particular values the corresponding attributes will have in a solution (Reitman, 1965; Voss & Post, 1988).

The lack of information at the beginning of the process and the complexity of ill-structured problems means that the representation phase that precedes the solution process becomes very important (Voss & Post, 1988). Ill-structured problems require extensive problem structuring before problem solving can begin (Simon, 1973; Goel & Pirolli, 1992). Such a problem structuring phase results in the construction of an initial problem space. In addition to the set of constraints that the problem solver spontaneously thinks of, an additional number of constraints will be expressed as the result of the consultation of external sources of information, e.g. books and experts. Enlarging the *information span*, i.e. the collection of relevant information, will provide the basis for formulating explicit constraints, and this will eventually result in a reduction of the number of acceptable solutions. In architectural design, the activity of structuring a problem is commonly referred to as the “analysis” stage. In this stage, the lack of information has to be overcome by exploring available information (Jones, 1963; Lawson, 1990).

The problem solving task then is to find a solution that satisfies all of a set of constraints (Simon, 1970). For ill-structured problems in general there are no strong

methods or algorithms that lead directly to a solution. The design problem is divided into sub-problems rather than solving the whole problem at once (Simon, 1973). Solving a sub-problem often requires specific information related to the particular sub-problem at hand. This information has to be acquired by searching through available information. Since there are no right or wrong answers a designer will use “personalized stopping rules” for deciding whether a solution satisfies constraints (Goel & Pirolli, 1992). The level of satisfaction thus depends on personal characteristics of the designer. According to Hamel (1990), an architectural designer can set an aspiration level for each constraint. If the aspiration level is easily attained the designer can aspire higher levels, if a solution is not found on a certain aspiration level, designers can lower their aspiration levels for satisfying. The characteristics of the solution will thus depend on the size and content of the set of constraints that are expressed in an individual problem space. A solution will be a good one to the degree that constraints are satisfied, and the information acquired from external sources influences the final design proposal through the set of constraints considered by an individual problem solver.

As a conclusion we can state that the design process forms a task that is “information intensive”. Information plays a different role depending on design phase, and both exploration and finding information are important sub-processes of the design process.

### 3. The empirical study

The analysis in the preceding section showed that in line with the description by Goel and Pirolli (1992) architectural design problem solving is a problem solving task in which the need for external information is high. When considering information systems for use before or during the design process two main purposes have been identified. First, in the problem structuring phase, the information span, i.e. the set of issues considered important, has to be enlarged. Second, during the later phases of the design process the finding of specific information is relevant when sub-problems have to be solved. In the design process, these two purposes are associated with two different kinds of information gathering behaviour: respectively *browsing* and *searching* (de Vries, van Andel & de Jong, 1992; de Vries, 1994). While the success of a browse session would be to increase one’s knowledge in a domain, the success of a search session would be to find relevant information on a particular topic within a domain. The present study was designed to determine whether there are advantages of certain information structures when carrying out a particular activity. For reasons of experimental control, specific tasks were introduced in order to bring on browse and search behaviour, rather than letting subjects free to use an information system during the design of an object (in our case a child play area). These tasks were identified in an exploratory study involving observation of subjects’ use of an information system during the design process (de Vries, 1994). A control task was introduced where subjects were asked to focus on surface characteristics of the information system (e.g. counting the number of figures in the pictures in the information system). The control task was used for establishing a baseline, since subjects in this condition were supposed to only superficially process the information during task performance. The experiment also introduced the

structure of the information as an independent variable (in our case information from environmental psychology on child play areas). First, a network structure using cross-references permitted to navigate from section to section. Second, a hierarchical structure was created with essentially the same information. In the hierarchy, the information was grouped into topics and sub-topics, and navigation took place by choosing topics and subtopics in order to reach a section. Finally, a mixed version of the system combined the two ways of navigation. A mixed system with the flexibility of traversing network nodes and the guidance of hierarchical structure is thought to combine the advantages of both approaches for its users (Girill & Luk, 1992).

### 3.1. EXPECTATIONS

The criterion tests chosen in this study correspond to the desired outcome in the context of a design task situation. For browsing, this consists in measuring the enlargement of information span. For searching, this consists in measuring the knowledge of specific topics. Finally, the count criterion test involved estimation of surface characteristics. In the experiment, subjects in all three task conditions were given all three criterion tests. A number of expectations were formulated regarding each of the dependent measures as follows.

(1) An *enlargement of the information span* is expected for browsing in all three structures. However, browsing the network is expected to manifest the largest increase as a result of the navigation possibilities offered by the cross-references. In the hierarchy, it is possible to navigate systematically through all branches, but the variety of topics addressed may be limited as a result of the organization into sub-topics. An enlargement of the information span is also expected as a result of searching depending on the number and variety of topics searched for. However, searching in the network can lead to seeing more information irrelevant to the current search topic than searching in the hierarchy condition, and therefore the enlargement in the network condition may be slightly larger than in the hierarchy condition. Subjects in the count control condition will pay much less or almost no attention to content and consequently are not expected to enlarge their information span.

(2) *Knowledge of particular topics* (the search topics) is expected to be high as a result of searching. Knowledge of the particular topics may also be reasonably high as a result of browsing, whereas counting is again expected to show lowest performance.

(3) *Estimation of surface characteristics*, e.g. the number of figures in the pictures shown, is expected to be fairly accurate as a result of counting, but not as a result of browsing or searching. Whereas the hierarchy and mixed condition permit a strategy where all branches of the hierarchical tree are systematically visited, the network condition does not permit such a strategy. Therefore, estimation of surface characteristics in the count condition should be more accurate in the hierarchy and mixed condition. Estimation of surface characteristics after searching or browsing may depend on how well the subjects understood the structure of the system, and on subjects' assumptions about the degree of coverage of the information in the system. These kind of estimations may also be important for reading strategies (McKnight, Dillon & Richardson, 1990).

### 3.2. METHOD

The general subject area involved in the study was the design of child play areas. All subjects completed a pre-test in which they were instructed to imagine receiving the assignment of the design of a child play area. After the pre-test, subject received their experimental assignments in the same context. The information system contained social science information on the relation between the environment and children's play behaviour.

#### 3.2.1. Subjects

Subjects were 3rd to 5th year students in architecture. These students have had some experience with design problems throughout their earlier years of study. Subjects were invited individually to come to the department and were paid for their participation. In total, 92 subjects participated in the experiment.

#### 3.2.2. The information system

The information in the system came from environmental psychology and emphasised the relationships between environment and behaviour in child play areas. The information on child play areas (adapted from Cohen, Hill, Lane, McGinty & Moore, 1979) was arranged into 56 sections. A section can be seen as an information unit or node in the system. Each section comprised of four elements: the title, the issue, the principle, and a recommendation. The sections are identified by their title. The issue explained shortly the main problem of the section. The principle provided a general solution to the problem. Finally, the recommendation was a more concrete idea of a solution in reality. This recommendation was also illustrated in a picture. In the system, each of the 56 sections occupied one screen.

Three variations on the linking structure of the information were created, as follows.

- (1) *Network*. The original references of Cohen *et al.* (1979) were used to create a network of information. Upon entering the system a starting screen appeared which showed nine entrances to the network. The subject could proceed by clicking on an entrance which caused a section to appear with cross-references at the bottom of the screen. The references at the bottom of the section permitted jumping to another related section and so on. These links can be considered relational (Mohageg, 1992).
- (2) *Hierarchy*. In another version of the system, the sections were clustered into topic groups. The starting screen in this version showed the two labels of the first sub-topics of the hierarchy. The sections could be attained by successively choosing sub-topics on four levels. This is a more organizational use of linking. No cross-references were given in this condition.
- (3) *Mixed*. A third version combined the two structures. The starting screen in this version showed both the two labels of the sub-topics in the hierarchy and the nine entrances to the network. Subjects were free to use either the hierarchical sub-topics or the network references. Regardless of the way in which a section was arrived at the subject could proceed by choosing a reference or by going up one level in the hierarchy.

In all three conditions, the subject could decide to go back to the starting screen by clicking on the word "Starting screen", or go back only one screen previously



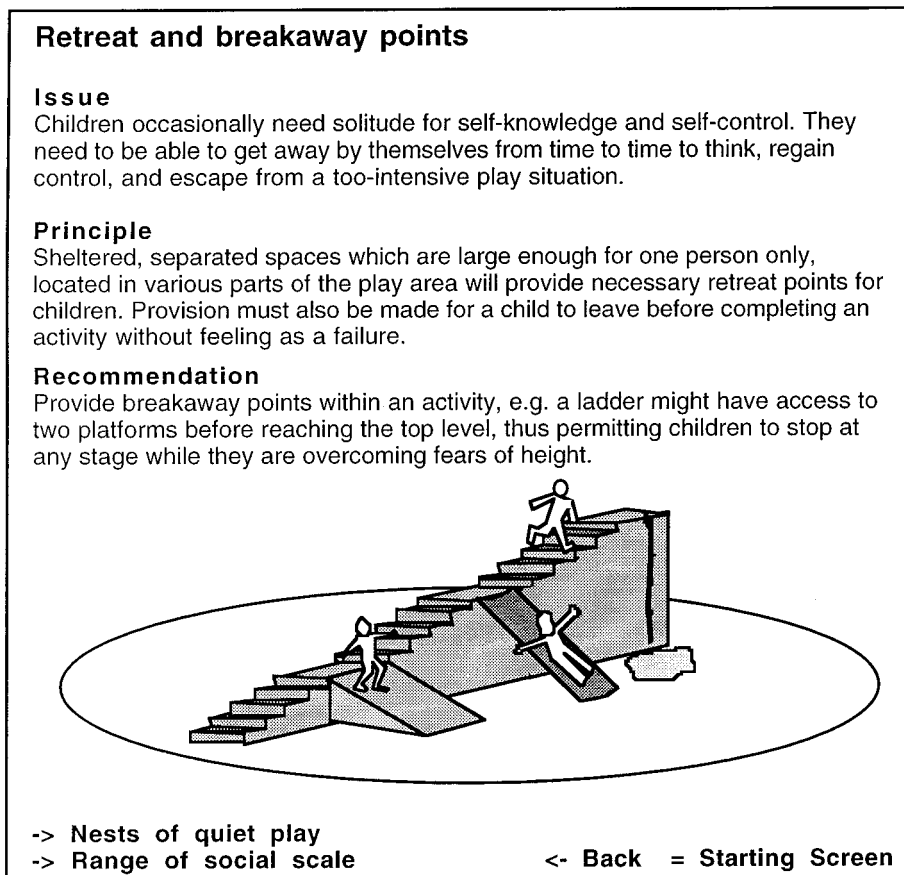


FIGURE 1. Example of a section with cross-references (network condition).

seen by clicking on the word “Back”. The names of the sections visited and visit times were written to a logfile.

### 3.2.3. Procedure

The experiment was held in a classroom, where up to 12 subjects could participate at the same time. First a pre-test was administered to assess the size of the information span prior to the experiment. Subjects were asked to write down their knowledge on the design of child play areas. They were instructed to imagine receiving the assignment of the design of a child play area near a school. Some characteristics of the school, such as number of children, were given in a short description. Subjects were invited to write down all the ideas s/he could think of regarding the design of such a child play area. It was explicitly stated that not only design objects could be named but that all kinds of design ideas were valid. Subjects were given 10 min to write down as many ideas as they could think of. Most subjects were already writing their reports within this period.

After the pre-test, subjects read the instructions about how to use the information

system. These instructions were different for the network, hierarchical, and mixed condition. Hereafter, they read the written task assignment. Three different task assignments were used as follows.

(1) *Browse*. The subjects in this condition were told to get to know as much as possible about designing child play areas. Subjects were told that they need not try to see everything in the system, they could choose topics of their interest. It was stressed that at the end of the computer session they should be able to name important aspects in the design of child play areas.

(2) *Search*. Subjects in this condition were told that they would get a number of topics on which they should search some information in the system. The subject was told to look for information on each topic. When they felt they had seen enough, or that nothing could be found, they could proceed by taking the next topic. They were given a list of 12 topics (e.g. storage, orientation) typed on individual sheets of paper and subjects searched the topics in the order in which they were given. They were told that at the end they should be able to report on the information found in the system.

(3) *Count*. This condition was introduced as a control task. Subjects were instructed to count the number of figures in the system as well as the number of arrows (pointing to the right). Subjects were told that at the end of the session they would be asked to give a reasonable estimation of the number of figures and arrows in the system.

Subjects were given 20 min to work with the system regardless of structure or assignment. They were told to work individually and they were not allowed to take notes.

The post-test consisted of criterion tests for all three tasks: browse, search, and count. For the browse criterion test, as for the pre-test, subjects were asked to write down as much ideas as possible. The search criterion test consisted of the 12 topics and the subjects were asked whether they found information about the topics and to write down what they knew about each topic. For the count criterion test, subjects were asked to estimate the number of figures and arrows in the entire system. In addition, they were asked to estimate the number of sections in the system for two reasons. Estimation of the number of sections in an information system is considered important in relation to the assumptions a subject has about the degree of coverage of the information in the system and for reading strategies (McKnight, Dillon & Richardson, 1990). Also, subjects are likely to estimate the number of figures and arrows evaluating the number of sections in the system. For example, a rough estimation of the number of figures can be made by multiplying the estimation of the number of sections with an estimation of the mean number of figures per section.

The order in which the tests were administered differed over task conditions. The browse criterion test preceded the search criterion test to avoid an influence of the given search topics on the information span. Furthermore, the criterion test corresponding to the particular task the subject had performed always came second. Three orders resulted: in the browse condition the order was count, browse, search criterion test. The subjects in the search condition received the criterion test in the order: browse, search and count. Finally, the subjects in the count condition

TABLE 1  
*Overview of the experimental design*

Pre-test	Task	Structure	<i>n</i>	Order of post-tests
Browse test	Browse	Network	10	Count-browse-search
		Hierarchy	11	
		Mixed	10	
	Search	Network	10	Browse-search-count
		Hierarchy	11	
		Mixed	9	
	Count	Network	10	Browse-count-search
		Hierarchy	10	
		Mixed	11	

answered first to the browse criterion test and then the count and search criterion test. Table 1 shows an overview of the experimental design.

#### 3.2.4. Analysis of the data

*Browse criterion test.* The pre-test and the post-test (browse) consisted of reports in which the subjects indicated the objects and ideas they would think of when designing a child play area. The statements in these reports had to be coded using a classification scheme. A classification scheme was developed to assess the information span in terms of the number of issues activated and the level of abstraction at which they were expressed. The reports were scored using meaningful units (Ferguson-Hessler & de Jong, 1990). A meaningful unit is a part of the report that can be defined by one class from the classification scheme. This means that the classes of the scheme have to be sharply defined. A unit can be a word, a part of a sentence or an entire sentence depending on the content and the context in which it occurs.

The classification scheme consisted of a list of issues (see Figure 2). An issue can be seen as an aspect of the designed object which will be important for a solution to be a qualitatively good solution. Notably, an issue can lead to the formulation of constraints. The issues were stated in very general terms, examples are safety, movement, circulation. Issues were derived from two sources: the reports from the pre-test and the information in the system. The first source, the reports from the pre-test, was examined to discover some of the frequently occurring issues. The second source, the information in the system, provided a list of issues that the subjects were supposed to acquire during task performance. These issues were thought to be important for scoring the post-test.

Statements may differ considerably in the level of abstraction in expressing an issue. In scoring the reports three levels of abstraction were distinguished: abstract concepts, performance requirements, and materializations. The highest level of abstraction consists of naming a concept like in “you should consider the *safety* of the children”. A more concrete statement would be “children should not cut themselves”. Such a statement is considered a performance requirement: it states a desired behaviour of the designed object. Finally, a statement containing reference

1. General architectural topics		
1.1 Neighbourhood level	1.1.1	Neighbourhood-based play
	1.1.2	A variety of possibilities in a neighbourhood
	1.1.3	Play areas as the centre of a neighbourhood
1.2 Play area level	1.2.1	Arrangement
	1.2.2	Circulation
	1.2.3	Accessibility
	1.2.4	Microclimate
	1.2.5	Groundcovering and berms
	1.2.6	Planting
	1.2.7	Basic utilities
	1.2.8	Maintenance
	1.2.9	Safety
	1.2.10	Use of materials and colours
2. Topics from environmental psychology		
2.1 Psychology	2.1.1	Perception, orientation
	2.1.2	Challenge
	2.1.3	Emotion
	2.1.4	Attraction
2.2 Activities	2.2.1	Rest
	2.2.2	Fantasy
	2.3.3	Creativity
	2.3.4	Natural elements (water, sand, fire)
	2.3.5	Education on environment (gardens, animals)
	2.3.6	Body movement—horizontal (running, etc)
	2.3.7	Body movement—vertical (climbing, etc)
2.3 Social topics	2.3.1	Different ages
	2.3.2	Different group sizes
	2.3.3	Views to and from play area, supervision

FIGURE 2. Classification scheme for the topics in the information span.

to materialization is considered the most concrete level of expressing an issue, e.g. naming an object or a property in reality as in “it should have a sand play area” or “make round edges of objects”. A report may be a list of abstract concepts like in “security, suitable for different ages, rest is important”, or a report may be a list of objects like in “fence, entrance, climbing object, slide, sand and water, trees” or a combination of these.

The levels were scored as 1, 2 and 3 respectively. If more statements were made following a certain statement but still concerning that statement, codes would be repeated (see Figure 3). This was done to account for writing style. For example, the sentence “*for safety I would design a fence and use rubber tiles*” should receive the same codes as “*for safety I would design a fence and for safety I would use rubber tiles*” (Safety: 1 3 1 3). In addition, a sentence like “*for safety I would design a fence and for body movement I would make a high object*” would get codes (Safety: 1 3, body movement: 1 3). All three statements are considered having the same number of elements with a similar degree of abstraction. The latter statement differs only in the number of different issues involved (two instead of only one).

A number of reports were scored separately by two raters. The codes were then

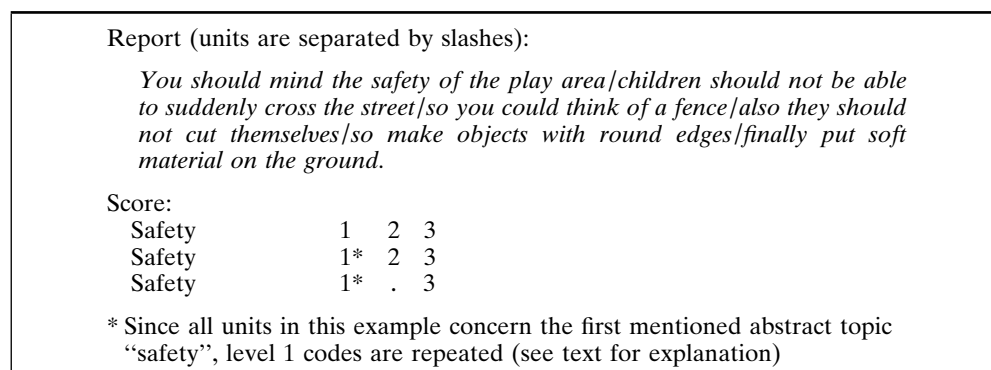


FIGURE 3. Scoring example.

compared, where the codes given were different the statement was discussed and the decision was made. Reliability was calculated on the basis of 10 post-test reports scored separately. The codes scored by the two raters were compared in two steps: issue and abstraction level. The percentage of agreement on the issues was 78% or 160 out of 205 lines in 10 reports. A line means a decision about a statement or a group of up to three statements if they belong together, cf. the number of lines of the report in Figure 3 equals 3. Cohen's *Kappa* was computed to correct for chance level, *Kappa* = 0.77. Reliability for the level of abstraction was calculated looking at the lines for which agreement existed on the issue. Percentage of agreement for the level of abstraction was 76% or 167 out of 220 codes (a line could consist of 1 to 3 codes). Cohen's *Kappa* was 0.66. The remaining reports (pre and post-test) were scored by one rater, and checked by the other rater, differences were noted and discussed until reaching agreement between raters. This resulted in only a few changes.

The size of the information span was calculated using the number of elements (over all three levels of abstraction) and the variety of issues addressed was measured by taking the number of different issues mentioned. The enlargement of the information span could be calculated by comparing the size and the variety in the pre-test and post-test. In addition the number of new issues in the post-test was measured, that is the number of issues that did not appear in the pre-test.

*Search criterion test.* The data of the search criterion test consisted of the subjects' answers to 12 topics. First, a list was made which contained relevant statements about each topic. These statements were derived from the information in the system. In the next step, the statements reported by each subject on each topic were scored by two persons. A statement could either be information from the system relevant to the topic or information from the system but irrelevant to the topic. For example, for the topic "storage" relevant statements were "close to play area" and "size adapted to children's height". First, the number of topics for which subjects produced at least one relevant statement was counted. Second, the number of relevant statements of each subject on each topic was counted. Statements from the information system but irrelevant to the topic were discarded.

*Count criterion test.* In the count criterion test, subjects were asked to estimate the number of figures, arrows and sections in the entire system. In each of the three experimental conditions the actual number of figures and the number of sections were the same (56 sections, 219 figures). The actual number of arrows varied for each of the three systems (225 in the network condition, 80 in the hierarchy condition and 305 in the mixed condition). In order to make comparisons between structure conditions the accuracy of the estimates was calculated as the percentage deviation from the actual number for each question as displayed in Equation 1.

$$\% \text{ deviation} = \frac{|Estimation - Actual \text{ number}|}{Actual \text{ number}} \times 100\% \quad (1)$$

### 3.3. RESULTS

The results of this study will be presented for each of the dependent variables separately. First, the students' performance at the pre-test is examined.

#### 3.3.1. Pre-test

On the pre-test, subjects were asked to write down ideas regarding the design of a play ground. With the help of the classification scheme, the size of the information span and the variety of issues within the information span can be assessed. The size of the information span was calculated by taking the total number of elements on the three abstraction levels together. As an indicator of the variety of issues in the information span the number of different issues was counted. Mean size of the information span for the entire sample was 18.7 elements (standard deviation 7.3) and mean variety was 7.9 different issues (standard deviation 2.4). The means for size and variety are displayed in Table 2.

In order to detect differences between conditions, a multivariate analysis of variance was carried out on size and variety of the information span with structure (network, hierarchy, mixed) and task (browse, search, count) as main between subjects factors. The analysis of variance showed no multivariate and no univariate effects of task and structure nor an interaction. It may therefore be concluded that the experimental groups did not differ in prior knowledge of the domain of designing child play areas.

#### 3.3.2. Browse criterion test

The post-test contained a browse criterion test analogous to the pre-test for prior knowledge. Table 3 shows the enlargement of the information span in size and variety expressed in the difference between pre-test and post-test. In addition, the number of new issues that did not appear in the pre-test, is displayed. From Table 3 it can be seen that the mean number of newly appeared issues is larger than the increase in variety of issues in all conditions. This means that besides naming new issues, at the same time issues named in the pre-test disappeared. While the increase in variety of issues gives an indication of the enlargement of the information span in terms of the number of themes underlying the information span, the number of new issues shows the degree to which new themes are adopted from the information in the system.

Multivariate analysis of variance showed an overall constant effect indicating that

TABLE 2  
*Prior knowledge: mean size and variety of the information span as a function of structure and task*

Task	Structure		
	Network	Hierarchy	Mixed
<b>Size</b>			
Browse			
<i>M</i>	16.7	20.3	18.3
<i>S.D.</i>	6.8	8.0	7.4
Search			
<i>M</i>	19.2	14.0	22.0
<i>S.D.</i>	5.7	9.0	6.6
Count			
<i>M</i>	20.8	18.3	19.3
<i>S.D.</i>	8.6	5.7	6.4
<b>Variety</b>			
Browse			
<i>M</i>	7.8	7.8	9.0
<i>S.D.</i>	2.3	2.4	2.4
Search			
<i>M</i>	8.0	6.2	8.4
<i>S.D.</i>	1.6	2.3	2.1
Count			
<i>M</i>	8.6	7.9	7.6
<i>S.D.</i>	2.6	3.1	2.2

an enlargement of the information span took place ( $F(3, 81) = 346.10$ ,  $p < 0.001$ ). Inspection of the univariate results revealed that this enlargement was significant for size, variety, and the number of new issues ( $F(1, 83) = 12.37$ ,  $p < 0.001$ ,  $F(1, 83) = 50.39$ ,  $p < 0.001$ ,  $F(1, 83) = 581.83$ ,  $p < 0.001$  respectively). The information span did benefit from working with the information system resulting in longer reports, more different issues, and a significant number of new issues.

The task main effect was significant ( $F(6, 164) = 3.12$ ,  $p < 0.01$ ) on the multivariate test. Again, the effect was consistent for all three dependent variables: size ( $F(2, 83) = 5.59$ ,  $p < 0.01$ ), variety ( $F(2, 83) = 3.56$ ,  $p < 0.05$ ) and new issues ( $F(2, 83) = 5.52$ ,  $p < 0.01$ ). The task main effect was tested making two comparisons. The browse and search conditions were compared with the count condition which acted as a control condition. Secondly, the browse condition was compared with the search condition. Examination of the individual parameters of these comparisons revealed that the task effect can be explained by a significant difference of the browse and search condition compared to the control count condition on all three dependent variables. The means in Table 3 show that as expected the enlargement was most important for the browse and search condition. The count condition shows minor pre-post-test differences in size and variety of the information span, but still a considerable amount of new issues. So, although subjects in the count condition did

TABLE 3  
*Browse criterion: enlargement (post-test–pre-test) and number of new issues in the information span as a function of structure and task*

Task	Structure		
	Network	Hierarchy	Mixed
<b>Enlargement (size)</b>			
Browse			
<i>M</i>	16.3	4.2	5.8
<i>S.D.</i>	11.8	9.6	8.6
Search			
<i>M</i>	1.6	3.8	3.7
<i>S.D.</i>	9.9	13.2	13.4
Count			
<i>M</i>	−1.4	0.4	0.6
<i>S.D.</i>	6.5	10.4	10.1
<b>Enlargement (variety)</b>			
Browse			
<i>M</i>	3.9	2.9	3.0
<i>S.D.</i>	3.5	4.8	3.0
Search			
<i>M</i>	2.4	3.5	3.3
<i>S.D.</i>	2.4	2.4	3.6
Count			
<i>M</i>	1.6	0.6	1.4
<i>S.D.</i>	3.3	3.6	3.3
<b>New issues</b>			
Browse			
<i>M</i>	7.3	6.6	6.7
<i>S.D.</i>	2.1	3.1	2.0
Search			
<i>M</i>	5.9	5.4	7.3
<i>S.D.</i>	2.2	1.9	3.2
Count			
<i>M</i>	5.1	4.6	5.0
<i>S.D.</i>	2.4	2.4	1.8

manage to grasp some new themes from the information in the system, their reports did not gain in length or variety.

The parameters for the second comparison, between the browse and search condition, was only significant as far as the enlargement in size was considered. Browsing resulted in a higher enlargement in size than searching.

The main effect for structure did not reach significance, nor did the interaction effect structure by task. In the introduction, we stated the expectation of an interaction between task and structure conditions. Especially, an interaction between browsing and searching in the network and hierarchy condition was expected because browsing in the network condition was thought to be the most beneficial to the size, variety and number of new issues in the information span. The



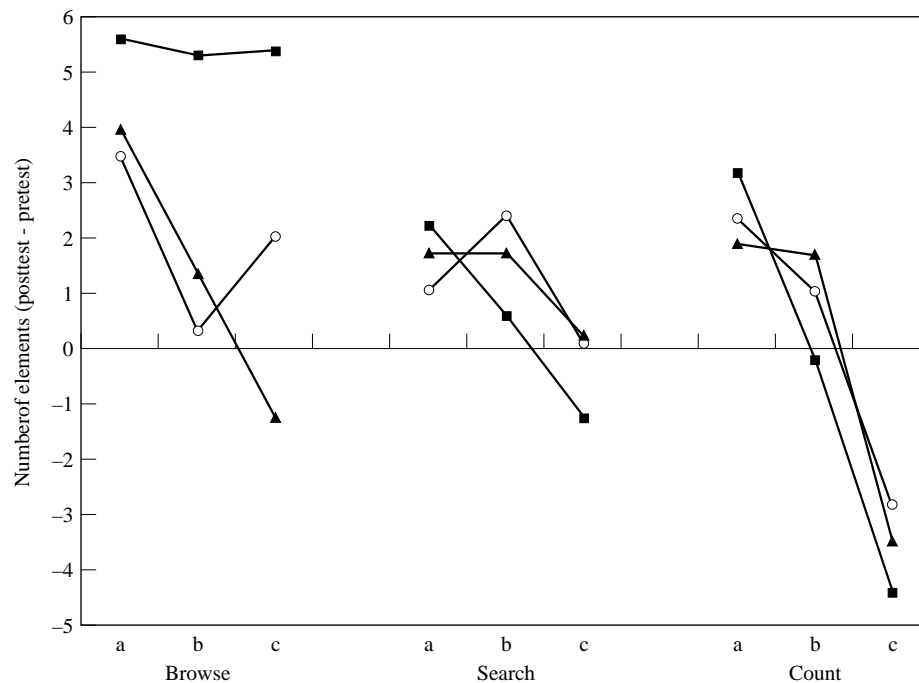


FIGURE 4. Changes in the number of elements in the information span as a function of structure, task and abstraction level (a = abstract concept, b = performance requirement, c = materialization). ■: network; ▲: hierarchy; ○: mixed.

parameter corresponding to this interaction therefore needs closer inspection. The analysis confirms an interaction in size, but not in variety and the number of new issues. When looking at the means in Table 3 it can be seen that the browse—network condition did indeed produce the most important enlargement in size. The means for variety and new issues, though not significant, show the same pattern.

In the preceding analyses, the size of the information span was measured taking the total number of elements in the subjects reports both before and after performing the experimental task. Figure 4 shows the changes in the number of elements as a function of the abstraction level at which the issues were expressed.

The figure shows that in general the number of abstract concepts, and performance requirements increases (post-test–pre-test difference  $>0$ ), whereas the number of materialization statements decreases. It should be noted that the pre-test reports contained more materializations, e.g. prototypical objects like sand play area, slide, etc. These objects were named to a lesser extent in the post-test. The opposite was true for abstract concepts. The number of abstract concepts increased as a result of using the system. The figure also shows that the large size increase in the browse–network condition, noted above, is almost equally distributed over all three abstraction levels.

### 3.3.3. Search criterion test

The search criterion test tested subjects' knowledge of specific topics. Subjects in the search condition are expected to answer to more topics on the search criterion test

TABLE 4  
*Search criterion: number of topics answered to and total search score as a function of structure and task*

Task	Structure		
	Network	Hierarchy	Mixed
<b>Topics answered to (maximum = 12)</b>			
Browse			
<i>M</i>	8.9	8.6	7.6
<i>S.D.</i>	1.3	1.6	1.6
Search			
<i>M</i>	8.6	8.0	9.1
<i>S.D.</i>	1.4	1.4	1.8
Count			
<i>M</i>	7.5	5.8	6.7
<i>S.D.</i>	2.4	2.4	2.1
<b>Search score</b>			
Browse			
<i>M</i>	18.2	16.3	15.3
<i>S.D.</i>	3.6	7.5	5.8
Search			
<i>M</i>	17.3	16.4	17.7
<i>S.D.</i>	4.2	6.0	4.1
Count			
<i>M</i>	11.3	10.3	12.0
<i>S.D.</i>	5.6	4.9	5.3

Three subjects from the search condition were left out because they only tried to answer to topics they actually searched for despite of the instructions to answer to all topics.

(maximum is 12) because they specifically searched information on these topics. In addition, they are expected to be able to tell more about each topic, resulting in a higher total search score. Table 4 shows the number of topics answered to with at least one relevant statement, and the search score (total number of relevant statements over all topics). Multivariate analysis of variance revealed a main effect of task ( $F(4, 160) = 5.27, p < 0.01$ ). This effect was significant for both the number of topics answered to ( $F(2, 80) = 9.54, p < 0.01$ ), and the search score  $F(2, 80) = 11.27, p < 0.01$ ). This effect can be attributed to the low performance of the subjects in the count condition. The subjects in the browse condition performed equally well compared to the subjects in the search condition.

A structure by task interaction should be found if the hierarchy and mixed condition are more suitable for searching. The analysis of variance did not show an interaction ( $F(8, 160) = 1.14, \text{n.s.}$ ). Table 4 shows that the means are comparable across structure conditions for each task. In addition, no main effect of structure was found ( $F(4, 160) = 0.88, \text{n.s.}$ ).

### 3.3.4. Count criterion test

The count criterion test consisted of estimation of surface characteristics, i.e. the number of figures and arrows in the information system. Since estimations

TABLE 5  
*Count criterion: mean % deviation, mean rank and number of subjects as a function of task*

Task	Mean	Rank	<i>n</i>
<b>% Deviation (number of sections)</b>			
Browse	63	43.84	31
Search	94	55.02	30
Count	67	40.92	31
<b>% Deviation (number of figures)</b>			
Browse	99	41.40	29
Search	86	56.76	29
Count	64	37.37	31
<b>% Deviation (number of links)</b>			
Browse	55	41.75	30
Search	110	56.96	28
Count	49	37.34	31

supposedly are related to the number of sections in the system, the estimation of the number of sections was also taken as a dependent variable. Some subjects gave very large estimations, but there is no reason to exclude them from analysis, since these estimations are valid observations given the task. An analysis of variance does not seem appropriate, because the large estimations have a great influence on the means and standard deviations. Therefore the Kruskal–Wallis test was used to test differences between the mean percentages of deviation from the actual numbers of sections, figures and links.

We expected that subjects in the hierarchy would make more accurate estimations. However, the Kruskal–Wallis test did not reveal a main effect for structure for any of the three estimations (sections, figures and links). Second, a main effect of task was expected. Subjects in the count condition should be able to make more accurate estimations. The Kruskal–Wallis test did indeed reveal a task main effect on the estimation of the number of figures ( $kw = 9.34, p < 0.01$ ) and the number of links ( $kw = 9.22, p < 0.01$ ) while the test for the number of sections only approached significance ( $kw = 4.77, p < 0.10$ ). Table 5 displays mean deviations from the actual values, mean rank per group and the number of observations per group. It can be seen from Table 5 that the subjects in the count condition indeed showed the lowest deviations from the actual numbers (the most accurate estimations) followed closely by the browse condition. The largest deviations from the actual numbers were found in the search condition as shown by the high mean rank in this group.

### 3.3.5. Exposure

Subjects in all nine conditions worked 20 min with the information system. Exposure to the information in the system is displayed in Table 6 expressed in the mean number of different sections visited.

The means in the table show that browsing in general seemed to result in higher

TABLE 6  
*Mean exposure (number of sections visited) as a function of  
 structure and task*

Task	Structure		
	Network	Hierarchy	Mixed
<b>Exposure</b>			
Browse			
<i>M</i>	35.1	29.9	29.1
<i>S.D.</i>	7.1	10.8	6.8
Search			
<i>M</i>	24.6	12.5	21.1
<i>S.D.</i>	3.8	2.4	2.8
Count			
<i>M</i>	30.8	24.8	26.6
<i>S.D.</i>	7.8	12.7	9.5

exposure. Furthermore, working with the network condition resulted in somewhat higher exposure as compared to the hierarchy and mixed conditions.

In combination with the three criterion tests, these data show that high exposure, such as when counting, does not necessarily lead to high information span, or to high knowledge of particular topics. Likewise, low exposure, such as when searching, is not necessarily associated to low information span, or to low knowledge of particular topics.

#### 4. Conclusions and discussion

Both tasks and variables used in this study were related to the characteristics of the design process as identified in the literature. The main purpose was to determine whether there are advantages of certain information structures when carrying out a particular activity.

As expected, an enlargement of the information span was found for browsing and searching, but not for counting. However, whereas browsing the network resulted in a large size enlargement, and browsing as well as searching the hierarchy in an intermediate size enlargement, only a small enlargement was found for searching the network. The latter result indicates that, in contrast to our expectations, for searching, a network structure in fact may not be useful for encountering meaningful additional information along the way. Nevertheless, a disadvantage for using the network for searching was not found either. The results on knowledge of particular topics showed that searching as well as browsing brought about relatively high scores in all structure conditions. These findings can partly be elucidated by appealing to the notion of divergent and convergent processes. Browsing is the more divergent process and can be influenced, i.e. encouraged or hindered, by information structure. Searching, in contrast, is the more convergent process that remains relatively unaffected by information structure. In de Vries (1994), some more evidence is presented in favour of this explanation. Finally, the criterion test of the count control condition involved estimation of surface characteristics, such as the

number of figures in the system. As expected, estimations were found to be most accurate as a result of counting, and least accurate for searching. However, the hierarchy and mixed structures did not seem to be particularly suitable for counting. Thus, the provision of organizational links does not necessarily render the resulting hierarchy and mixed structures transparent for something as elementary as apprehending the size of the system, even if this is the explicit aim of system use.

In conclusion, evidence was found in favour of some, but not all, of our expectations. Some interactions between information access and task characteristics are shown, but further research is needed to examine the suitability of candidate structures for specific purposes, in particular, by broadening the range of performance measures.

In the introduction, two methodological issues in the study of the relation between type of task and type of information access were identified: developing both appropriate experimental tasks and appropriate dependent measures of success. In this study, the use of hypertext information systems was investigated in context, as opposed to studies that either focus on delivery of a working prototype without worrying about user aspects, or focus on user aspects only as far as question answering is concerned. Both tasks and measures were derived in the context of the architectural design process. However, whereas in design problem solving different types of information gathering behaviour will occur spontaneously, time on system was kept constant in our study in order to systematically investigate the effects of task and structure. The advantage of such a more context-rich approach is the higher ecological validity of the tasks studied. Our tasks were of the type observed in real design problem solving and subjects' received an instruction evoking a design context. In these circumstances, the subjects' cognitive state (Waterworth & Chignell, 1991) during task performance is likely to match their condition during a real design setting.

A second advantage concerns the development of measures of success that are meaningful in a particular context. As obvious as it may seem, performance has to be linked to the goal of system use. Nowadays, systems are built that are claimed to support browsing as well as searching. Yet most research still largely concentrates on performance for question answering. Whereas so called browse questions are now subject of study, performance statements still refer to speed and accuracy, and at best, "completeness" of the answer (Waterworth & Chignell, 1991; Rada & Murphy, 1992: p. 11). Furthermore, an abundance of dependent variables exists for measuring navigation and orientation in hypertext (see for example in Canter, Rivers & Storrs, 1985; Edwards & Hardman, 1989; Simpson & McKnight, 1990). Despite their merits for describing how hypertext systems are used, these measures are not self-evident in the study of the use of hypertext systems for real-life tasks. In our study, the variable exploration (Canter *et al.*, 1985) was used in combination with three variables related to purposes in information gathering, and ultimately to their related outcomes. Studying navigational characteristics in combination with measures related to specific purposes could be a promising approach for research into hypertext use for a variety of tasks.

The last issue to be addressed here is the generalization of the findings. The subjects in our study were students with approximately the same amount of experience in design projects, and hence formed a relatively homogeneous group

with intermediate experience in designing. Such a homogeneous group of experienced architects is not available, neither in number nor in background. In addition, architects would have brought a large variety of their own experiences to the laboratory (see Rowland, 1992), which would have made it difficult to use them as subjects in this kind of experiment. Furthermore, large individual differences in information gathering strategies could have been found. However, such individual differences in spontaneous information gathering behaviour during the design task may be an important factor in the success of information systems. Information gathering varies with expertise, from novices through intermediates to experts. For example, Rowland (1992), in a comparison of novice and expert instructional designers, found that experts interpreted a problem as poorly defined by the given information and consequently devoted a lot of time to problem structuring and the exploration of information. Novices interpreted a problem as well defined by the given information and did little exploration. Therefore, generalizing from the students in our study to experts could mean an even larger accent on the problem structuring phase of the design process.

Finally, indications that problem solving is similar across design disciplines comes from studies encompassing more than one discipline, such as the study by Goel and Pirolli (1992) on architectural, mechanical and instructional design. This means that the results of our study can be applied to design tasks in general. Nevertheless, in order to generalize, it is important that the tasks possess certain essential characteristics, such as the need to consult external information sources during the process. In fact, research in other design disciplines has come up with recommendations for design information systems which resemble the network and mixed structures used in our study. For example, the provision of information structures that do not impose a hierarchical structure, but that show the complexity of information in design domains, is recommended for instructional design (Pirolli & Russell, 1990), software design (Guindon, 1990) and mechanical design (Visser, 1990). Future research should aim at clarifying the relation between task characteristics and information structures during design problem solving both within and across these disciplines.

This work was carried out at, and supported by, the Eindhoven University of Technology as part of the Ph.D. dissertation of the first author. We have appreciated the input of Joost van Anel (Eindhoven University of Technology) as an expert in the field of environmental psychology and the stimulating suggestions of Jules Pieters (University of Twente), who acted as the overall supervisor of the Ph.D. dissertation.

## References

- AKIN, Ö. (1986). *Psychology of Architectural Design*. London: Pion.
- BEGORAY, J. A. (1990). An introduction to hypermedia issues, systems and application areas. *International Journal of Man-Machine Studies*, **33**, 121-147.
- BRANSFORD, J. D., FRANKS, J. J., MORRIS, C. D. & STEIN, B. S. (1979). Some general constraints on learning and memory research. In L. S. CERMAK & F. I. M. CRAIK, Eds. *Levels of Processing in Human Memory*, pp. 331-354. Hillsdale, NJ: Lawrence Erlbaum.
- CANTER, D., RIVERS, R. & STORRS, G. (1985). Characterizing user navigation through complex data structures. *Behaviour and Information Technology*, **4**, 93-102.
- COHEN, U., HILL, A. B., LANE, C. G., MCGINTY, T. & MOORE, G. T. (1979)

- Recommendations for child play areas*. Technical Report R79-1, Centre for Architecture and Urban Planning, University of Wisconsin-Milwaukee, Milwaukee.
- CONKLIN, J. (1987). Hypertext: an introduction and survey. *IEEE Computer*, **9**, 17–41.
- DUMAIS, S. T. (1988). Textual information retrieval. In M. HELANDER Ed. *Handbook of Human-Computer Interaction*, pp. 673–700. Amsterdam: North Holland.
- EDWARDS, D. M. & HARDMAN, L. (1989). 'Lost in hyperspace': cognitive mapping and navigation in a hypertext environment. In R. McALEESE Ed. *Hypertext: Theory into Practice*, pp. 105–125. Oxford: Intellect Limited.
- FERGUSON-HESSLER, M. G. M. & DE JONG, T. (1990). Studying physics texts: differences in study processes between good and poor performers. *Cognition and Instruction*, **7**, 41–54.
- GIRILL, T. R. & LUK, C. H. (1992). Hierarchical search support for hypertext on-line documentation. *International Journal of Man-Machine Studies*, **36**, 571–585.
- GOEL, V. & PIROLI, P. (1992). The structure of design problem spaces. *Cognitive Science*, **16**, 395–429.
- GROSS, M. D., ERVIN, S. M., ANDERSON, J. A. & FLEISHER, A. (1988). Constraints: knowledge representation in design. *Design Studies*, **9**, 133–143.
- GUINDON, R. (1990). Designing the design process: exploiting opportunistic thoughts. *Human-Computer Interaction*, **5**, 305–344.
- HAMEL, R. (1990). *Over het denken van de architect: een cognitief psychologische beschrijving van het ontwerpproces bij architecten* [On designing by architects: a cognitive psychological description of the architectural design process]. Ph.D. dissertation, University of Amsterdam, Amsterdam.
- JONES, J. C. (1963). A method of systematic design. In J. C. JONES & D. G. THORNLEY Eds. *Conference on Design Methods*, pp. 53–74. London: Pergamon Press.
- LAWSON, B. (1990). How designers think, 2nd Edn. London: The Architectural Press.
- MARCHIONINI, G. & SHNEIDERMAN, B. (1988). Finding facts versus browsing knowledge in hypertext systems. *IEEE Computer*, **20**, 70–80.
- McKNIGHT, C., DILLON, A. & RICHARDSON, J. (1990). A comparison of linear and hypertext formats in information retrieval. In R. McALEESE & C. GREEN Eds. *Hypertext: state of the Art*, pp. 10–19. Oxford: Intellect Limited.
- MOHAGEG, M. F. (1992). The influence of hypertext linking structures on the efficiency of information retrieval. *Human Factors*, **34**, 351–367.
- NEWELL, A. & SIMON, H. A. (1972). *Human Problem Solving*. Englewood Cliffs NJ: Prentice Hall.
- PIROLI, P., & RUSSELL, D. M. (1990). The instructional design environment: technology to support design problem solving. *Instructional Science*, **19**, 121–144.
- RADA, R., & MURPHY, C. (1992) Searching versus browsing in hypertext. *Hypermedia*, **4**, 1–30.
- REITMAN, W. R. (1965). *Cognition and Thought*. New York: Wiley.
- ROWLAND, G. (1992). What do instructional designers actually do? An empirical investigation of expert practice. *Performance Improvement Quarterly*, **5**, 65–86.
- SIMON, H. A. (1970). Style in design. In J. ARCHEA & C. EASTMAN Eds. *EDRA TWO, Proceedings of the Second Annual Environmental Design Research Association Conference*, pp. 1–10. Pennsylvania: Dowden, Hutchinson & Ross, Inc.
- SIMON, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence*, **4**, 181–201.
- SIMPSON, A. & McKNIGHT, C. (1990). Navigation in hypertext: structural cues and mental maps. In R. McALEESE & C. GREEN Eds. *Hypertext: State of the Art*, pp. 73–83. Oxford: Intellect Limited.
- VISSER, W. (1990). More or less following a plan during design: opportunistic deviations in specification. *International Journal of Man-Machine Studies*, **33**, 247–278.
- VOSS, J. F. & POST, T. A. (1988). On the solving of ill-structured problems. In M. T. H. CHI, R. GLASER, & M. J. FARR, Eds. *The Nature of Expertise*, pp. 261–285. Hillsdale, NJ: Lawrence Erlbaum.
- VRIES, E. DE, (1994). *Structuring information for design problem solving*. Ph.D. dissertation, Eindhoven University of Technology, Eindhoven, The Netherlands.

- VRIES E. DE, VAN ANDEL, J. & DE JONG, T. (1992). *Computer-aided transfer of information on environment and behaviour: network versus hierarchical structures*. Paper presented at the 23th Annual Environmental Design Research Association Conference, Boulder, USA.
- WADE, J. W. (1977). *Architecture, Problems and Purposes; Architectural Design as a Basic Problem-Solving Process*. New York: John Wiley.
- WATERWORTH, J. A. & CHIGNELL, M. H. (1991). A model for information exploration. *Hypermedia*, **3**, 35–58.
- WRIGHT, P. & LICKORISH, A. (1990). An empirical comparison of two navigation systems for two hypertexts. In R. MCALEESE & C. GREEN, Eds. *Hypertext: State of the Art*, pp 84–93. Oxford: Intellect Limited.

Paper accepted for publication by Editor, B. R. Gaines.