Reading Patterns and Usability in Visualizations of Electronic Documents

KASPER HORNBÆK and ERIK FRØKJÆR
University of Copenhagen

We present an exploration of reading patterns and usability in visualizations of electronic documents. Twenty subjects wrote essays and answered questions about scientific documents using an overview+detail, a fisheye, and a linear interface. We study reading patterns by progression maps that visualize the progression of subjects' reading activity, and by visibility maps that show for how long different parts of the document are visible. The reading patterns help explain differences in usability between the interfaces and show how interfaces affect the way subjects read. With the overview+detail interface, subjects get higher grades for their essays. All but one of the subjects prefer this interface. With the fisheye interface, subjects use more time on gaining an overview of the document and less time on reading the details. Thus, they read the documents faster, but display lower incidental learning. We also show how subjects only briefly have visible the parts of the document that are not initially readable in the fisheye interface, even though they express a lack of trust in the algorithm underlying the fisheye interface. When answering questions, the overview is used for jumping directly to answers in the document and to already-visited parts of the document. However, subjects are slower at answering questions with the overview+detail interface. From the visualizations of the reading activity, we find that subjects using the overview+detail interface often explore the document further even when a satisfactory answer to the given question has already been read. Thus, overviews may grab subjects' attention and possibly distract them.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation/methodology, Interaction styles*; I.3.6 [**Computer Graphics**]: Methodology and Techniques—*Interaction techniques*

General Terms: Experimentation, human factors, measurement, performance

Additional Key Words and Phrases: Electronic documents, digital documents, information retrieval, information visualization, reading, reading patterns, overview+detail interface, fisheye interface

1. INTRODUCTION

Reading of electronic documents has become ubiquitous and deeply integrated in our everyday activities. Such documents are read on the World Wide Web, in

Authors' addresses: Kasper Hornbæk, Natural Sciences ICT Center, University of Copenhagen, Universitetsparken 5, DK-2100, Copenhagen, Denmark, email: KHornbaek@nik.ku.dk; Erik Frøkjær, Department of Computing, University of Copenhagen, Universitetsparken 1, DK-2100, Copenhagen, Denmark; email: erikf@diku.dk.

electronic journals, in professional work, and as part of recreational activities. Sellen and Harper [1997, 2002] describe the use of paper and electronic documents among analysts at the International Monetary Fund and assess that 14% of the time analysts worked with documents they used electronic documents only. Analysts used a combination of paper and electronic documents 35% of the time. Byrne et al. [1999] studied World Wide Web usage and found that users spend at least twice as much time using the information they find, compared to searching, browsing, or any other activity. In the study by Byrne et al., reading is the main activity in using information.

Unfortunately, users experience a variety of difficulties when reading electronic documents. These difficulties include cumbersome navigation [Dillon 1994; O'Hara and Sellen 1997], a lack of overview of the document [O'Hara and Sellen 1997], lower tangibility of electronic documents compared to paper [Hansen and Haas 1988], an unclear awareness of the length of documents [O'Hara and Sellen 1997], lower reading speed caused by the poor resolution of most screens [Mills and Weldon 1987; Dillon 1994], learning of lower quality compared to paper documents [Hertzum and Frøkjær 1996], and possible fatigue if reading for extended periods of time.

As a potential solution to these problems and with the aim of improving the ubiquitous reading activity, visualization techniques have been used for presenting electronic documents [Eick et al. 1992; Hornbæk and Frøkjær 2001]. Some visualizations of electronic documents show the contents of a document together with an overview of that document [Eick et al. 1992; Graham 1999]. Others show a distorted version of the document compressed to fit a limited amount of screen space [Robertson and Mackinlay 1993] or consisting of only the important parts of the document [Furnas 1986; Kaugars 1998]. However, the usability of visualizations of electronic documents is largely unexamined and to our knowledge no one has investigated if such interfaces change how users read.

In this paper, we analyze how visualization techniques support reading of electronic documents. We compare a linear, a fisheye, and an overview+detail interface used in an experiment by 20 subjects for writing essays and answering questions about scientific documents. We use logged data about the interaction process to visualize subjects' reading activity. Our visualizations help describe reading patterns by showing how reading progresses and for how long certain parts of a document are visible. The reading patterns give insight into how the interfaces affect subjects' reading activity and into how we can design interfaces that better support reading. In addition, we investigate the common hypothesis that overview+detail and fisheye interfaces improve usability. Extending our previous analysis [Hornbæk and Frøkjær 2001], we use the reading patterns to explain differences in usability between the interfaces.

In the next section, we outline previous work on visualization and on studies of reading patterns in electronic documents. Section 3 describes our experiment on visualizing electronic documents. Section 4 describes the reading patterns. Section 5 presents the differences in usability between interfaces and explains them with reference to the reading patterns. In Section 6 the results are discussed and Section 7 presents our main conclusions.

2. BACKGROUND

2.1 Visualization of Electronic Documents

Visualizations of electronic documents are of two kinds: overview+detail and distortion-based interfaces. Overview detail interfaces show an overview of the document separated from the detailed content [Plaisant et al. 1995]. The overviews show zoomed-out representations of the document [Eick et al. 1992; Boguraev et al. 1998; Graham 1999] or thumbnail representations of the pages in the document [Adobe Acrobat[1]; Ginsburg et al. 1996]. On some overviews, occurrences of query terms in the document are color-coded [Graham 1999; Byrd 1999]. Besides the present paper, we know of no evaluations of overview+detail interfaces for electronic documents. However, Chen and Rada's [1996] review of research in hypertext suggests that overviews improve users' effectiveness. Studies of text overviews also suggest improved performance from having an overview of an electronic document. Studies of Superbook [Egan et al. 1989] compared the performance of subjects who used a 562-page paper manual for a statistics package to subjects searching an electronic version of the manual using an expandable table of contents (i.e. a text overview) combined with the detailed contents of the manuals. With the most developed version of Superbook, 10 subjects performed 25% better than subjects searching in the paper manual. In two experiments, Dee-Lucas and Larkin [1995] compared linear text to overview interfaces in which the overview and the detailed contents were not visible simultaneously. When reading an approximately 2000-word physics text, the subjects using the overview had better and broader recall of text topics compared to subjects without the overview.

Distortion-based interfaces show the entire document in a limited amount of screen space or show only the most important parts of the document. Robertson and Mackinlay [1993] proposed an interface that shows only one part of a document in focus and the other pages of the document zoomed out to fit the remaining space. Holmquist [1997] describes a similar interface that can use semantic zooming on the pages that are out of focus. In other distortion-based interfaces, only important parts of the document are readable. Importance may be determined by structural properties of the document, such as sections and subsections [Páez et al. 1996]; by the current view of the document [Furnas 1986]; or by similarity between the terms used for retrieving the document and the sections of the document [Kaugars 1998]. Páez et al. [1996] describe a zoomable user interface for electronic documents where title, headings, and key sentences are larger than other parts of the document. Initially, the entire document is visible on the screen. When comparing this interface to a hypertext interface, Páez et al. [1996] found no difference between interfaces in 36 subjects' satisfaction, task completion time, or memory for the contents of the document.

[1]http://www.adobe.com/products/acrobat

2.2 Reading Patterns

A large literature describes how interface designs, tasks, genre characteristics, and reader traits influence performance when reading electronic documents [Wright 1987; Hansen and Haas 1988; Dillon 1994; Muter 1996; Schriver 1997]. Here we focus on characterizing patterns in reading activity—how readers navigate and manipulate documents as they try to accomplish their aims with reading. Three kinds of reading patterns are discussed in the literature.

In one reading pattern, documents are read in a non-linear fashion, occasionally with multiple readings of some sections. Bazerman [1988, pp. 235–253] discusses how the purposes and background knowledge of seven physicists influence the way they read academic papers. In general, papers were read selectively with jumps between different sections. Readers often looked for new information or for particular sections, such as the method section in descriptions of empirical research. In addition, parts of the documents were given multiple readings at different intensity. Dillon [1994, pp. 93–101] describes two series of 15 interviews about how participants read academic papers and software manuals. For academic papers, most readers skim titles and author names, after which they scan the abstract and main sections. Then, important sections are read non-linearly or the whole paper is read serially. In software manuals, the participants most often consulted the table of contents or the index sections to get a feel for the contents and locate useful places for reading. Horney and Anderson-Inman [1994] describe the reading patterns of 17 middle school students in two hypertext stories. From logged interaction with the stories, they identify different processes in the reading activity such as skimming, checking, reading, responding, studying, and reviewing. Horney and Anderson-Inman also show how students read the stories multiple times and how students sometimes read the story from end to beginning.

As a second reading pattern, linear reading occurs under some circumstances. Goldman and Saul [1990] showed that the most common reading strategy among students reading informational texts was to read linearly through the text once. Foltz [1996] compared the reading strategies in two hypertexts and a linear document. Independently of document type and task type (reading for general knowledge vs. reading for finding specific information), 80 to 90% of the transitions to new sections and pages were coherent with the overall organization of the text. In a second experiment, Foltz used verbal reports to show that when subjects answered specific questions they read linearly from text preceding the desired information and towards that information, apparently trying to maintain the coherence of the text. Similarly, subjects in the experiment of Hertzum et al. [2001] often begin reading sections preceding the section containing the answer to the question posed. Seemingly, subjects try to establish the context of the answer.

A third group of reading patterns is formed by the various roles played in reading by different parts of a document. For academic papers, certain sections—those containing dense formulas or problem formulations—might be skipped entirely [Bazerman 1988]. Bishop [1999] used focus groups and interviews to investigate how readers of scientific papers use document components.

She shows how readers use document components, such as the abstract or figures, for orientation, for gaining an overview of the paper, for directing attention, for comprehension, and for inspiring additional reading. In addition, readers often jump non-linearly between different parts of the paper.

In summary, reading patterns are diverse and no one has studied reading patterns for visualizations of electronic documents. The investigations of reading patterns in overview+detail and distorted interfaces, and the techniques employed to describe the reading patterns are new in this study.

3. EXPERIMENT

To investigate how visualizations of electronic documents influence reading patterns, we conducted an experiment where subjects answered questions and wrote essays about documents in a domain of their interest. Subjects completed these tasks using a linear, a fisheye, and an overview+detail interface. Below we describe the interfaces and the experiment; Hornbæk and Frøkjær [2001] contained a preliminary account of the usability data from the experiment but only a brief mention of reading patterns, our main focus here.

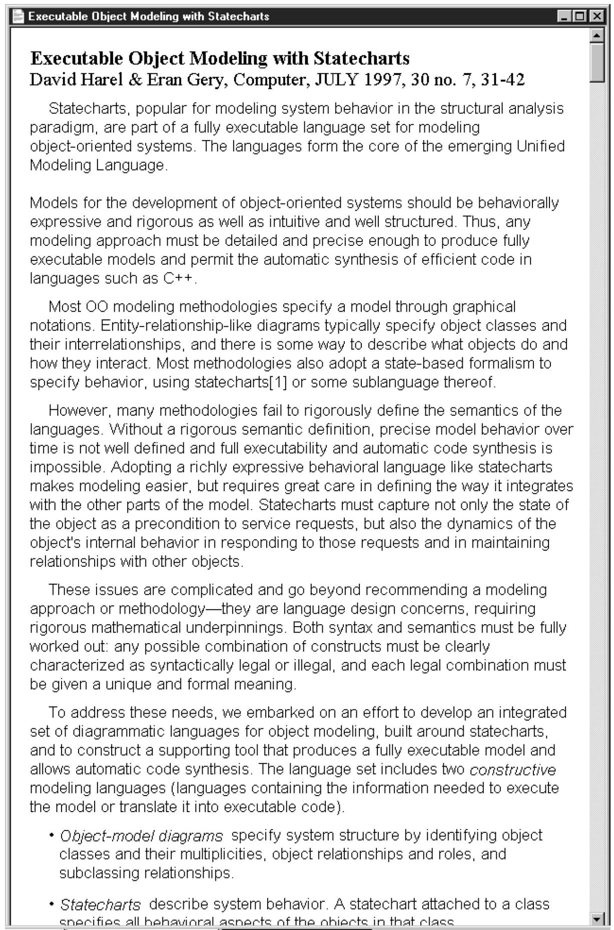
Our experiment is exploratory, aimed at describing reading patterns and how interfaces affect reading. In addition, we had two hypotheses about differences between interfaces.

1. Based on the literature described in Section 2 we expected the overview+detail interface to improve satisfaction and task completion time over the linear interface. We expected this because the overview+detail interface facilitates navigation by providing the overview pane and because this interface presents the reader with an overview of the structure and contents of the entire document.
2. We also expected the fisheye interface to decrease task completion time because the documents are compressed in the presentation and therefore less time-consuming to navigate. The fisheye interface was also expected to support readers in employing an overview-oriented reading style, so-called outlining [Anderson and Armbruster 1982]. One measurable implication of this reading style is faster reading, since subjects quickly establish an overview of the text.

3.1 Interfaces

We compared a linear, a fisheye, and an overview+detail interface. Figures 1(a), 1(b), and 1(c) show the three interfaces. In these interfaces, documents can be navigated using the mouse or the keyboard. Subjects may *highlight* words in the documents. By entering one or more words in a dialog box, all instances of the entered words are highlighted in red in the document. From a pop-up menu, the highlights can be removed.

In the linear interface, Figure 1(a), the document is shown as a linear sequence of text and pictures. This interface is similar to most interfaces in practical use and serves as a baseline against which the other interfaces can be compared.



Executable Object Modeling with Statecharts

Executable Object Modeling with Statecharts
David Harel & Eran Gery, Computer, JULY 1997, 30 no. 7, 31-42

Statecharts, popular for modeling system behavior in the structural analysis paradigm, are part of a fully executable language set for modeling object-oriented systems. The languages form the core of the emerging Unified Modeling Language.

Models for the development of object-oriented systems should be behaviorally expressive and rigorous as well as intuitive and well structured. Thus, any modeling approach must be detailed and precise enough to produce fully executable models and permit the automatic synthesis of efficient code in languages such as C++.

Most OO modeling methodologies specify a model through graphical notations. Entity-relationship-like diagrams typically specify object classes and their interrelationships, and there is some way to describe what objects do and how they interact. Most methodologies also adopt a state-based formalism to specify behavior, using statecharts[1] or some sublanguage thereof.

However, many methodologies fail to rigorously define the semantics of the languages. Without a rigorous semantic definition, precise model behavior over time is not well defined and full executability and automatic code synthesis is impossible. Adopting a richly expressive behavioral language like statecharts makes modeling easier, but requires great care in defining the way it integrates with the other parts of the model. Statecharts must capture not only the state of the object as a precondition to service requests, but also the dynamics of the object's internal behavior in responding to those requests and in maintaining relationships with other objects.

These issues are complicated and go beyond recommending a modeling approach or methodology—they are language design concerns, requiring rigorous mathematical underpinnings. Both syntax and semantics must be fully worked out: any possible combination of constructs must be clearly characterized as syntactically legal or illegal, and each legal combination must be given a unique and formal meaning.

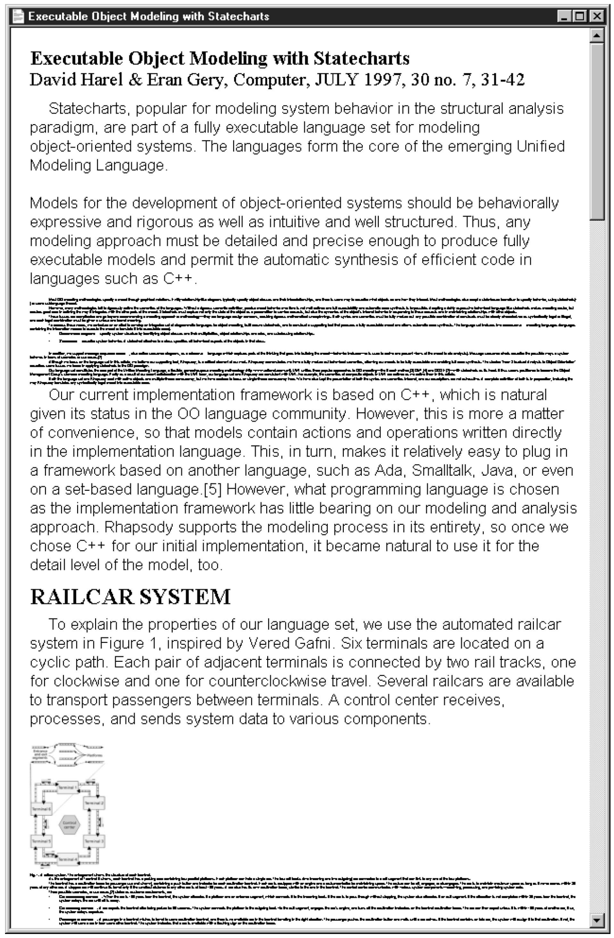
To address these needs, we embarked on an effort to develop an integrated set of diagrammatic languages for object modeling, built around statecharts, and to construct a supporting tool that produces a fully executable model and allows automatic code synthesis. The language set includes two *constructive* modeling languages (languages containing the information needed to execute the model or translate it into executable code).

- *Object-model diagrams* specify system structure by identifying object classes and their multiplicities, object relationships and roles, and subclassing relationships.

- *Statecharts* describe system behavior. A statechart attached to a class

Fig. 1(a). The linear interface.

In the fisheye interface, Figure 1(b), certain parts of the document are considered more important than other parts. The most important parts of a document are always readable. The other parts are initially distorted below readable size, but can be *expanded* and made readable if the user clicks on them with the mouse. To *collapse* a part of the document back to its distorted state, the user simply clicks on that part again. When over a part of the document, the mouse cursor changes to indicate whether that part can be expanded, collapsed, or neither of the two (as is the case for example for the beginning of a document). Because of the distortion, the initial size of the documents in the fisheye interface was on average 25% of their sizes in the linear interface. Two strategies are used for determining which sections are important. First, sentences selected from the beginning and end of a document unit are among the best indicators of the contents of that unit [Bradow et al. 1995; Kupiec et al. 1995]. Therefore, the first and last paragraphs of a section are considered important. This scheme is recursively applied to subsections, so that when a section is expanded only



Executable Object Modeling with Statecharts

Executable Object Modeling with Statecharts
David Harel & Eran Gery, Computer, JULY 1997, 30 no. 7, 31-42

Statecharts, popular for modeling system behavior in the structural analysis paradigm, are part of a fully executable language set for modeling object-oriented systems. The languages form the core of the emerging Unified Modeling Language.

Models for the development of object-oriented systems should be behaviorally expressive and rigorous as well as intuitive and well structured. Thus, any modeling approach must be detailed and precise enough to produce fully executable models and permit the automatic synthesis of efficient code in languages such as C++.

Our current implementation framework is based on C++, which is natural given its status in the OO language community. However, this is more a matter of convenience, so that models contain actions and operations written directly in the implementation language. This, in turn, makes it relatively easy to plug in a framework based on another language, such as Ada, Smalltalk, Java, or even on a set-based language.[5] However, what programming language is chosen as the implementation framework has little bearing on our modeling and analysis approach. Rhapsody supports the modeling process in its entirety, so once we chose C++ for our initial implementation, it became natural to use it for the detail level of the model, too.

RAILCAR SYSTEM

To explain the properties of our language set, we use the automated railcar system in Figure 1, inspired by Vered Gafni. Six terminals are located on a cyclic path. Each pair of adjacent terminals is connected by two rail tracks, one for clockwise and one for counterclockwise travel. Several railcars are available to transport passengers between terminals. A control center receives, processes, and sends system data to various components.

Fig. 1(b). The fisheye interface.

the first and last parts of the subsections are immediately readable. Second, as mentioned in Section 2, readers often attend to and find certain components of a document especially useful [Dillon 1994; Bishop 1999]. Therefore, abstracts and section headings are always visible, and graphics and tables are diminished less than text.

In the overview+detail interface, Figure 1(c), the document is shown as a linear sequence of text and pictures (the detail pane) together with a tightly coupled overview of the document (the overview pane). The documents used in the experiment were on average 17 times longer than the amount of text visible at one time in the detail pane. The ratio between the overview of the document and the entire document (i.e. the magnification or zoom ratio) was therefore on average 1:17. A rectangular field-of-view covering a part of the overview pane indicates which part of the document is currently shown in the detail pane. The field-of-view can be moved to change which part of the document is shown in the detail pane. On the overview pane, section and subsection headings are shown at a

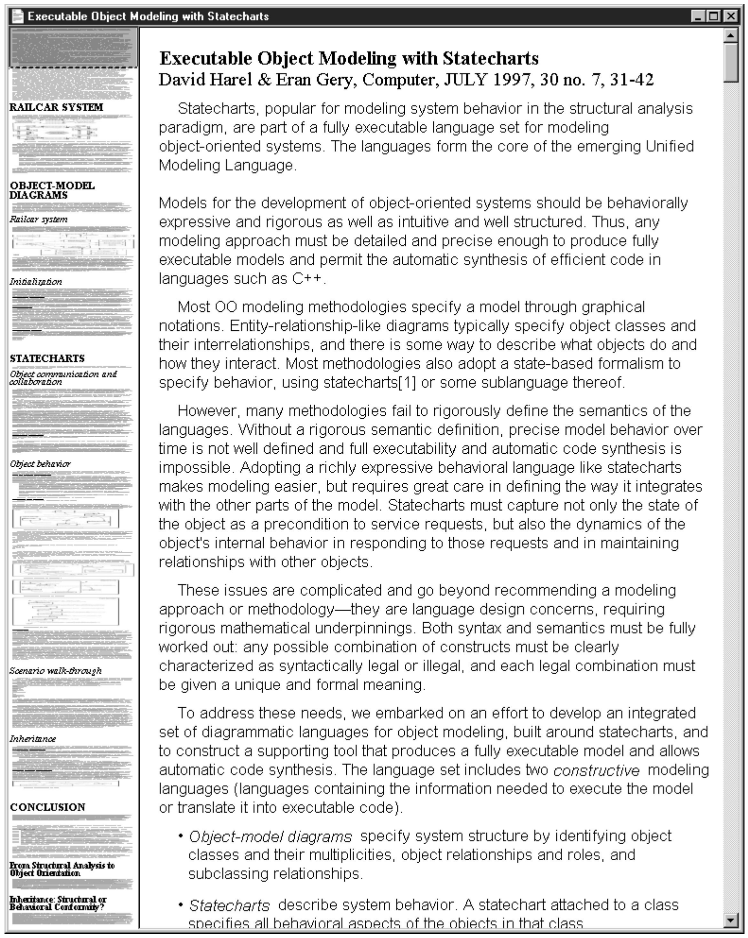


Fig. 1(c). The overview+detail interface.

fixed, readable size. Except for the headings, the contents of a section are shrunk to fit the remaining space allocated to show that section. We believe that the readability of headings and the stability of the overview pane are the main improvement over previous overview+detail interfaces for electronic documents, for example Graham [1999]. Our implementation scales to arbitrarily large documents, as long as the number of headings to be shown on the overview is kept relatively constant. However, the practical utility of the overview+detail interface probably declines above a 1:25 ratio between overview size and document length, which is the recommended ratio in other contexts [Plaisant et al. 1995].

Note that the overview+detail interface uses more screen real estate than the two other interfaces. We designed the interfaces this way to ensure that the normal/expanded content of the document is displayed in an identical way between interfaces (compare Figures 1(a), (b), and (c)). Superficially, an overview+detail interface the same size as the linear interface would make for a cleaner comparison. However, then either the line length or the font size would have been different between the interfaces—and both line length and font size are known to impact reading speed (see for example Hansen and Haas [1988]; Marcus [1992]; Dyson and Haselgrove [2001]). Therefore, we chose to design the interfaces so that the areas for displaying document content had similar sizes. A practical argument for our design of the overview+detail interface is that for interfaces for electronic documents horizontal space is often not fully used because of the aspect ratio of most screens and the recommendation that relatively short lines make for optimal reading speed. In contexts where single documents are read, horizontal space is thus often available.

3.2 Tasks and Documents

Subjects were given two types of tasks: essay tasks and question-answering tasks. The essay tasks and the question-answering tasks correspond to reading to understand a document and reading to answer a question. These aims of reading are central in several accounts of typical reading tasks, for example Schriver [1997]. Although answering questions is obviously a typical task with electronic documents, it may be argued that no one reads an entire document from the screen. However, our intention with the overview+detail and the fisheye interface is to make online reading more attractive and thus we need to look at tasks that make subjects read to understand.

In essay tasks, subjects read a document to learn its main contents. Afterwards and without access to the document, they were required to write a one-page essay, stating the main theses and ideas of the document, and one page of personal comments about the document. After writing the essays, subjects were given six incidental-learning questions. An example of an incidental-learning question is: 'Which integrity problems can occur in what the author calls the simple business application architecture?'

In question-answering tasks, subjects were required to answer six questions about a document, one question at a time. The six questions were varied as to (1) position in the document where the answer can be found (in the first or last part of the document), (2) how easily accessible the sentences or sections containing the answer are (whether they are near section beginnings, tables or figures), and (3) the usefulness of the words of the question as terms for highlighting (whether or not the question contained terms that were located near the answer). An example of a question is: 'What is, according to the paper, the biggest problem in relation to automatically transforming procedural code to object-oriented code?'

The documents used in the experiment were six IEEE journal papers from the Digital Library Initiative test bed at University of Illinois at Urbana-Champaign [Bishop 1995]. All documents were on topics within object-oriented systems development. The paper versions of the documents were between 8 and 14 pages long. The documents contained figures, tables, formulas, and text. From our presence during the experiment we conclude that no subjects had previously read any of the papers. To achieve uninterrupted reading and increased realism, we did not impose a time limit on the tasks. However, subjects were made aware of how much time they had used when reading one paper for more than one hour, or when they took more than 30 minutes to answer one of the six questions about a document. The descriptions of the tasks, the answers to the tasks, the training material, and the satisfaction questionnaires were all in the native language of the subjects, Danish.

3.3 Subjects

The subjects in the experiment were students at the Department of Computing, University of Copenhagen, who had chosen to participate in a course involving the experiment. The subjects had studied computer science for a mean time of 6.5 years. Of the 20 subjects, 15 were males and five females, with a mean age of 27. Sixteen subjects reported using computers every day, four subjects several times a week. Fourteen subjects reported familiarity with object-oriented systems development from courses, 11 subjects had such familiarity from systems development projects.

3.4 Design

The experiment employed a within-subjects factorial design, with the independent variables being interface type (linear vs. fisheye vs. overview+detail) and task type (essay vs. questions-answering). The experiment consisted of three sessions. In each session the 20 subjects used one interface to solve a task of each type. Each session lasted approximately one hour and 45 minutes, giving a total of 106 hours of experimental data. Tasks and interfaces were systematically varied and counterbalanced. We formed six groups based on permutations of the three interfaces. Using L to designate the linear, F the fisheye and O+D the overview+detail interface, these groups used the following orders of interfaces: L•F•O+D, L•O+D•F, F•L•O+D, F•O+D•L, O+D•L•F, O+D•F•L. Because six was not a divisor of the number of subjects, four groups comprised three subjects and two groups comprised four subjects. The tasks for these six groups were found by randomly choosing Latin squares such that the three interfaces and the three sessions had an approximately equal number of different tasks.

3.5 Reading Patterns and Usability Measures

One of the contributions of this paper is the description of reading patterns based on visualizations of reading activity described in Section 4. Reading patterns are described in terms of reading modes and events. To ensure that modes and events were reliably detected, one of the authors first developed a classification of reading modes and events and applied it on all visualizations of reading activity. The classification was performed blind to which interface the subjects had used. Afterwards, the other author classified a random sample of 20% of the visualizations of reading activity. The Pearson correlations between the authors' estimation of the duration of reading modes in essay tasks were between .96 and .99. For reading events, only the classification of one task differed. For question-answering tasks, the correlations were between .89 and .97. For the analysis in section 4, we used the classification of visualizations in the sample agreed upon by the authors. Visualizations not in the sample were adjusted to reflect the consensus among the authors.

To uncover the usability of the interfaces, we measured the following:

—*Grades* were given to all tasks. The answers were graded blind by the first author—without any knowledge of which subject had made the answer or with which interface the answer had been made. We used a five point grading scale, ranging from zero—a missing or completely wrong answer—to four—an outstanding and well-substantiated answer. For the question-answering tasks, grades were given according to how many aspects of the question the answer covered. A classification of the main ideas in the documents and important aspects of questions was developed to assist a systematic and uniform grading.

—*Incidental learning* was measured as the number of correct answers to incidental learning questions, resulting in a score from 0 to 6.

—*Task completion time* was used as the indicator of efficiency. All subjects' interactions with the interfaces were logged and the task completion times were derived from the logged data. For essay tasks, only the time spent reading is considered task completion time, leaving out the time spent writing the essay.

—*Satisfaction* was measured in three ways. After using each interface, subjects answered twelve questions about the perceived usability of the interface and their experiences with solving the tasks. After having used all three interfaces, subjects indicated which they preferred. Subjects also wrote comments about the interfaces after using each of them, and described why they preferred using one of the interfaces.

3.6 Procedure

The experiment took place in a lab without external disturbances. Two subjects participated at a time. Upon arriving, subjects filled out a questionnaire on background information and on their familiarity with object-oriented systems development. Then, subjects were trained until they felt confident in operating the interfaces. Training was supported by a two-page description of how to operate the interfaces. The subjects also completed three training tasks, which introduced them to the interfaces, and the question-answering and essay tasks. The mean time used to complete the training tasks was 35 minutes. After training, the subjects completed the first session of the experiment. Subjects returned the next day to the lab and completed the remaining two sessions.

The subjects received the tasks on sheets of paper, on which they also wrote the answers for the question-answering tasks. After finishing reading documents, the subjects proceeded immediately to the writing of essays, for which they received paper and pencil. The subjects were not allowed to take notes while reading the documents.

3.7 Analysis

The experimental design was expected to result in 20*3 solutions to the essay tasks, but one subject did not complete a task, and one solution was dropped from the analysis because of a time usage three interquartile ranges above the

75-quartile, leaving 58 solutions. The task completion time for that solution was 163 minutes, in comparison to the overall average of 42 minutes. For the question-answering tasks, the design should give 360 (20*3*6) answers, but one subject failed to complete a task, leaving 354 answers.

We analyzed the data by ANOVAs with interface type, task, session, and subject as factors; interfaces were compared using linear contrasts [Rosenthal et al. 2000]. Essay tasks and question-answering tasks were analyzed separately.

4. READING PATTERNS

4.1 Reading in Essay Tasks

4.1.1 *Progression Maps and Reading Modes.* We visualize each subject's reading activity for an essay task using what we call progression maps. The progression maps show what parts of a document subjects can see at which time in the reading process. Figure 2 shows an example of a progression map for an essay task. On the progression maps, we identified three modes to describe how subjects read a document (see Figure 2). In the *initial orientation mode*, subjects navigated through the document in a non-linear fashion. We found this mode at the beginning of a task, if the subject attempted initial orientation. The initial orientation mode ends when subjects began reading linearly through the document from the beginning. In the *linear read-through mode,* subjects read through the document from the beginning to the end in a linear way, with occasional skips forwards and backwards. This mode ended when subjects began to navigate non-linearly through the document for more than one minute and did not return to continue the linear read-through. In the *review mode,* subjects looked again at what they presumably felt were the most important sections in the document in a non-linear order. This mode was found at the end of a task.

In every task, we found a linear read-through. In 34 tasks we found an initial orientation mode and in 56 tasks a review mode. To show the individuality and variety in reading patterns, Figure 3 gives four examples of progression maps.

Figure 4 shows the average duration of the three reading modes. We found significant differences in time spent in the modes for the initial orientation mode ($F[2, 32] = 3.38$, $p < .05$). In the fisheye interface, more time was spent in the initial orientation mode ($M = 4.6$ min., $SD = 5.5$) compared to the linear ($M = 2.1$ min., $SD = 3.2$, $F[1, 32] = 5.02$, $p < .05$) and the overview+detail interface ($M = 2.0$ min., $SD = 3.2$, $F[1, 32] = 5.11$, $p < .05$). A significant difference between the interfaces was also found in the time spent in the linear read-through mode, $F[2, 32] = 10.86$, $p < .001$. A linear contrast shows that subjects spend only two-thirds as long with the fisheye interface in the read-through mode ($M = 26.6$ min., $SD = 16.2$) as with to the other two interfaces (linear: $M = 37.0$ min., $SD = 10.6$, $F[1, 32] = 15.23$, $p < .001$; overview+detail: $M = 37.5$ min., $SD = 11.7$, $F[1, 32] = 17.25$, $p < .001$). For the review mode, we find no significant difference, $F[2, 32] = 1.48$, $p > .2$.

We made two further observations about the reading behaviour in the initial orientation mode. First, on the progression maps we repeatedly

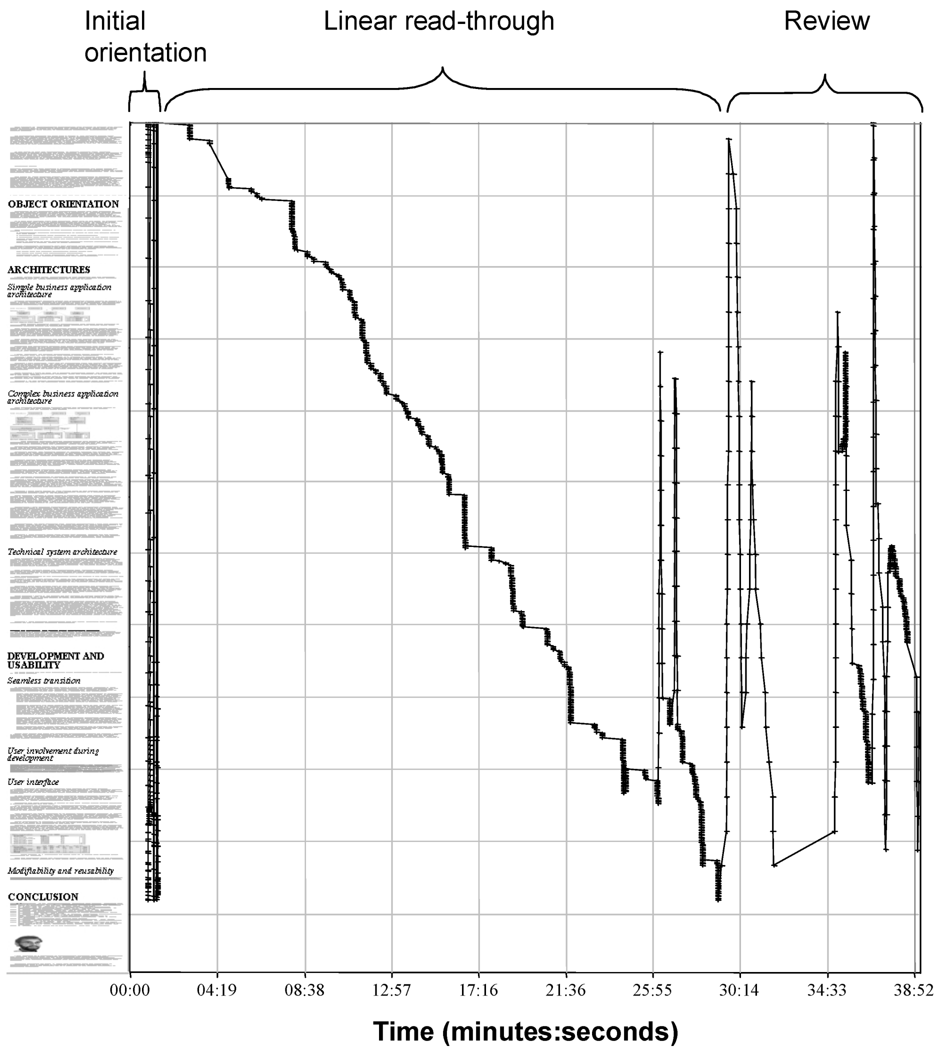


Fig. 2. Progression map showing reading modes. This figure shows a progression map for a subject doing an essay task. The reading modes are indicated at the top of the figure. The horizontal axis shows time elapsed since the beginning of the task. The vertical axis shows the position in the document visible to the subject as the top-most position in the detail window. The vertical axis also shows an overview of the contents of the document. In the figure the grid size indicates the height of one screen in the linear and overview+detail viewer. For the fisheye interface, subjects can see approximately twice as much.

observed an orienting behaviour from the subjects that we call *flip-through*. In a flip-through, subjects scrolled through the entire document in less than 30 seconds (see Figure 2 for an example). Subjects did so at the beginning of an essay task. This behaviour seems similar to flipping through the pages in a book or a journal. We observed flip-throughs in 30 out of the 59 essay tasks, with no difference between interfaces. Subjects may have used flip-throughs for obtaining an overview of the documents, a task

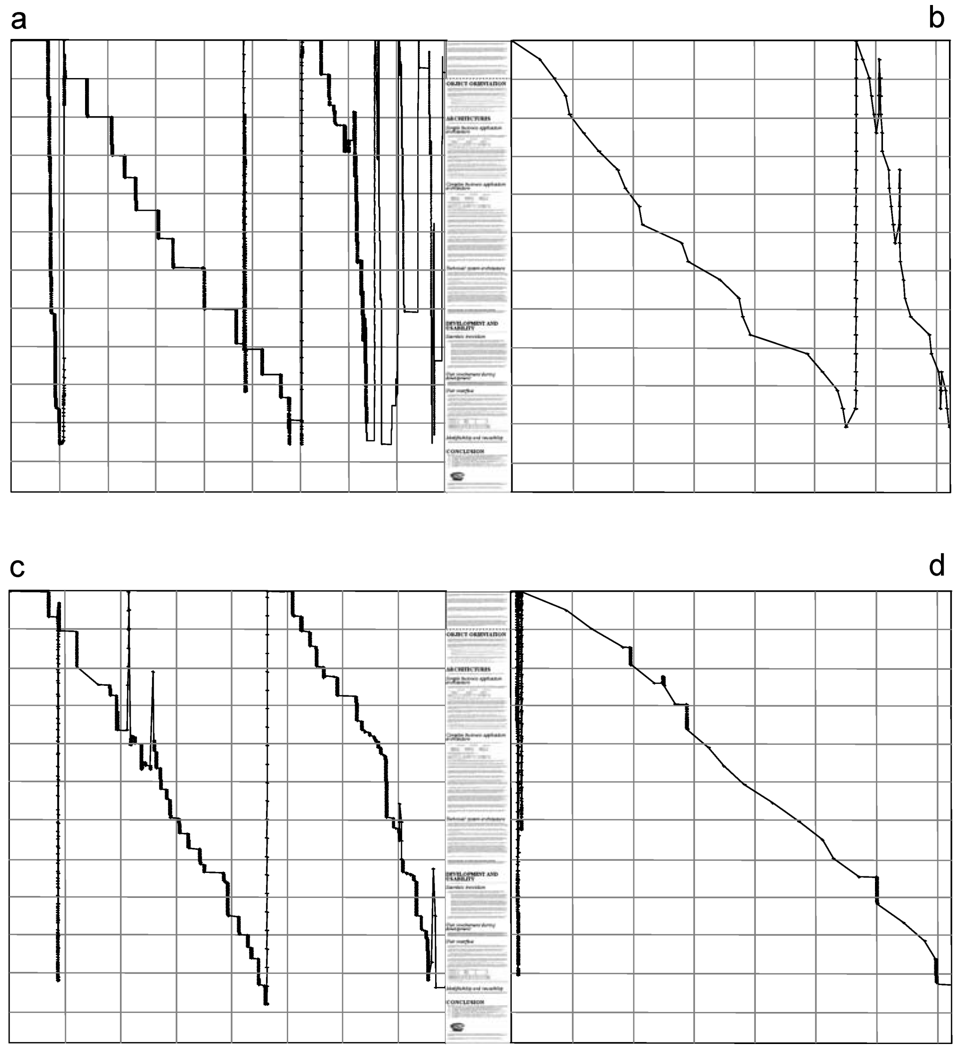


Fig. 3. Examples of reading patterns for the same essay task. Panel **a** shows a subject who has a brief initial orientation (including a flip-through), a linear read-through mode (including a quick digression back to the beginning of the document), and a review mode. Panel **b** shows a subject who immediately begins reading linearly through the document, and who in the review mode goes from the beginning to the end of the paper. Panel **c** shows a subject who has a lengthy read-through phase (including something that looks like a flip-through) and a long review phase, which almost seems to be a re-read of the document. Panel **d** shows a subject with a brief initial orientation phase (comprised of a flip-through) and a lengthy linear read-through. The task completion times are approximately **a** 39 min., **b** 62 min., **c** 45 min., and **d** 31 min.

that is notoriously difficult for electronic documents [O'Hara and Sellen 1997].

Second, we noticed that subjects during the initial orientation mode almost exclusively looked at the introduction and the conclusion of the paper, see Table I.

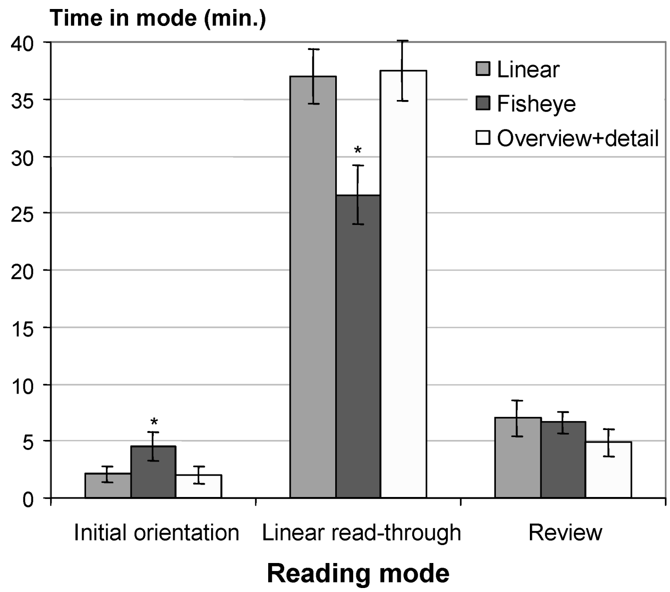


Fig. 4. Time spent in reading modes. The figure shows the average time subjects spent in the three reading modes for each of the interfaces. Subjects without a certain mode were counted as spending zero minutes in that mode. An asterisk denotes a significant difference between interfaces. Error bars show the standard error of the mean.

Table I. Document Parts Visible in the Initial Orientation Mode. The Table Shows the Percentage of the 34 Tasks with an Initial Orientation Mode where the Document Parts Described in the Left-Most Column are Visible for More Than One Minute

Document part	*Percentage tasks with parts visible*
Introduction and abstract	76%
Conclusion	41%
Other sections	18%
References and appendices	12%

4.1.2 *Expansion and Collapsing of Sections in the Fisheye Interface.* When using the fisheye interface, subjects on the average expanded 90% (SD = 18) of the sections in a document—see Figure 5. Six subjects in one or more tasks expanded all sections at once by selecting the pop-up menu item 'expand all'; the rest of the subjects expanded sections by clicking with the mouse on the section. We also examined in what reading modes subjects expanded sections or kept previously expanded sections expanded. Our hypothesis was that the fisheye interface should support an overview-oriented reading style, meaning that subjects expanded sections primarily in the linear read-through mode. In the initial exploration mode, subjects expanded or kept expanded approximately one fourth (M = 22%, SD = 32) of the sections in the document. In the linear read-through mode, subjects expanded or kept expanded 85% (SD = 24)

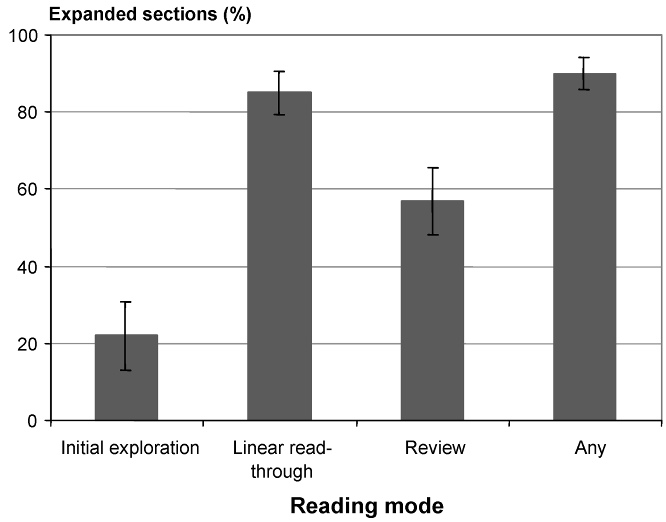


Fig. 5. Percentage expanded sections in each reading mode. This figure shows the average number of sections that subjects expanded or that were kept expanded in the three reading modes. The rightmost bar shows the number of sections open in any reading mode. Error bars show the standard error of the mean.

of the sections. In the review mode, subjects expanded or kept expanded approximately half (M = 57%, SD = 37) of the sections.

4.1.3 *Visibility Maps.* For all essay tasks, we also visualize reading activity by visibility maps. The maps were made by arbitrarily dividing the document into 100 parts of equal length. For each subject, we replayed the logged interaction and registered which parts were visible and for how long. Note that for the fisheye interface, only initially readable or expanded sections are counted as visible. Figure 6 shows an example of a visibility map for an essay task. To test the differences between interfaces revealed by casual inspection of the visibility maps, we compared the average percentage of the reading time spent in collapsed versus initially readable parts of the documents—see Figure 7. These maps, and the accompanying tests, reveal three interesting patterns about how long different parts of the documents were visible.

First, the visibility maps show that the relative duration for which different parts were visible differs between interfaces—see Figure 7. Here, we look at those parts of the documents that in the fisheye interface were initially readable. We found a significant difference between the interfaces in the duration these parts were visible, $F[2, 32] = 35.2$, $p < .001$ (we used the arcsine transformation on the percentage values before running ANOVAs, as differences in proportion violate assumptions underlying ordinary ANOVAs—see Cohen [1988]). In the fisheye-interface (M = 13%, SD = 4.4), the initially readable parts were visible for approximately 50% longer than in the linear interface (M = 8%, SD = 2.7, $F[1, 32] = 56.3$, $p < .001$) and the overview+detail interface (M = 9%, SD = 3.2, $F[1, 32] = 48.8$, $p < .001$). Similarly, we find a difference between interfaces in how long parts, which in the fisheye interface were initially collapsed,

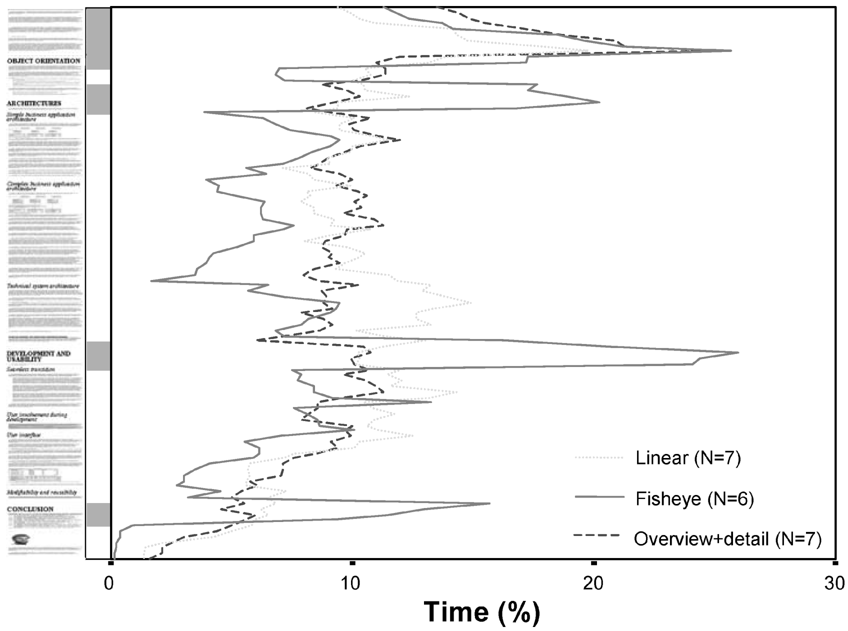


Fig. 6. Visibility map for one of the three essay tasks. The horizontal axis shows the average time a part of the document is visible for each interface. The vertical axis shows position in the document, as indicated by the overview of the document. The grey squares along the vertical axis indicate parts of the document that were initially readable in the fisheye interface.

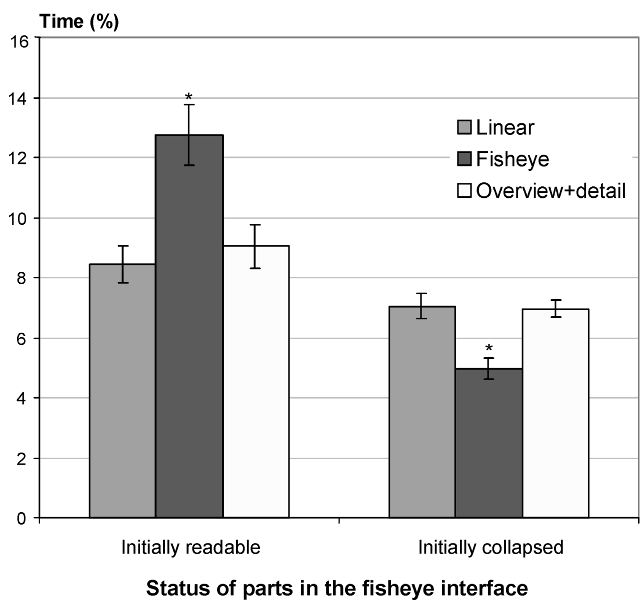


Fig. 7. Time spent in parts of the documents that are initially readable vs. initially collapsed in the fisheye interface. This figure shows the average time spent in those parts of the documents that in the fisheye interface are either initially readable or initially collapsed. An asterisk indicates a significant difference between the interfaces. Error bars show the standard error of the mean.

were visible, $F[2, 32] = 36.0$, $p < .001$. Linear contrasts show that in the fisheye interface ($M = 5\%$, $SD = 1.5$) these parts were visible for a shorter time compared to the other two interfaces (linear: $M = 7\%$, $SD = 1.9$, $F[1, 32] = 60.2$, $p < .001$, overview+detail: $M = 7\%$, $SD = 1.3$, $F[1, 32] = 46.9$, $p < .001$).

Second, in the overview+detail interface and the linear interface, subjects have sections visible a comparable length of time.

Third, for the linear and overview+detail interface we find a difference between how long certain parts of the document are visible compared to the fisheye interface. The time spent in parts of the document that in the fisheye interface are initially readable is longer with the overview+detail and linear interface compared to the time spent in parts that are initially collapsed in the fisheye interface. As the linear and overview+detail interfaces have comparable task completion times, this holds for both absolute time and relative to the total task completion time. Thus, our algorithm for the fisheye interface chooses sections to be initially readable that subjects in the other two interfaces spend a relatively long time reading.

4.2 Reading in Question-Answering Tasks

For question-answering tasks, we visualize reading activity for each subject's answer to each of the six questions on a progression map. To analyze these maps, we use a notion of *targets* in the documents, of reading events called *first contact*, and reading modes called *target reading* and *further explorations* (see Figure 8). A *target* is a part of the document in which an answer to the current question can be found. For each question-answering task, the answer to two of the six questions (or a substantial part of the answer) can be found in more than one place. The progression maps show the target as a point, but obviously text just before and after the target point is also important. Thus, we consider a subject to do *target reading* as long as the target is visible in the browser window and while the target is less than half a screen-length above the top of the detail window. For the fisheye interface, contents in the visible area may be collapsed, allowing targets further down the text to be read. For this interface, we therefore used one screen length above the top of the window to delimit the target area. Note that these measures are entirely based on text length and that no attempt to delimit when a target begins or ends has been made. Figure 9 shows individual reading patterns in question-answering tasks.

4.2.1 *First Contact with a Target and Direct Jumping.*

First contact is the moment when any target becomes visible for the first time in the detail part of the interface. To be considered a first contact, the target area must be visible for at least 20 seconds. If the target is at the beginning of the document, the subject might have been reading the task description. Therefore, we begin to look for a first contact after 10 seconds. Figure 10 shows the average time passed from the beginning of the task to the moment when subjects make first contact. We find no difference between interfaces in how fast subjects made first contact, $F[2, 313] = .341$, $p > .5$.

In some tasks, subjects went directly to a target by clicking on the overview pane, a *direct jump*. In 54 out of the 120 tasks solved with the overview+detail

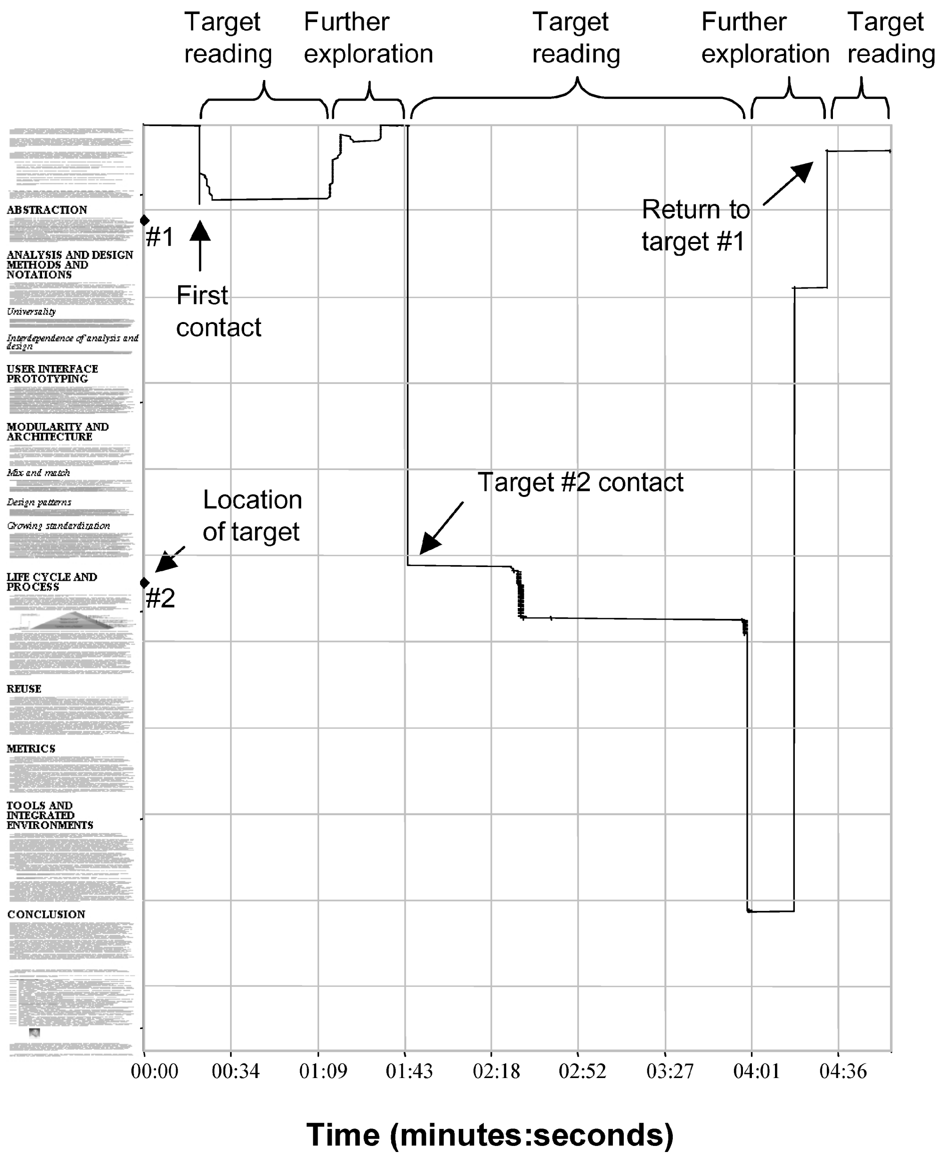


Fig. 8. An example of a progression map for one question in a question-answering task. The reading modes are indicated at the top of the figure. The horizontal axis shows time elapsed since the beginning of the task. The vertical axis shows the position in the document visible to the subject as the top-most position in the detail window. The vertical axis also shows an overview of the contents of the document. In addition, the two targets (#1 and #2) and the reading events are shown on the map. In the figure the grid size indicates the height of one screen in the linear and overview+detail viewer. For the fisheye interface, subjects can see approximately twice as much.

interface, subjects made first contact this way. Subjects therefore seem able to relate the information on the overview to the questions. In 13 tasks, subjects return to an already visited target by a direct jump. Since subjects only return in 44 of the tasks solved with the overview+detail interface, direct jumping

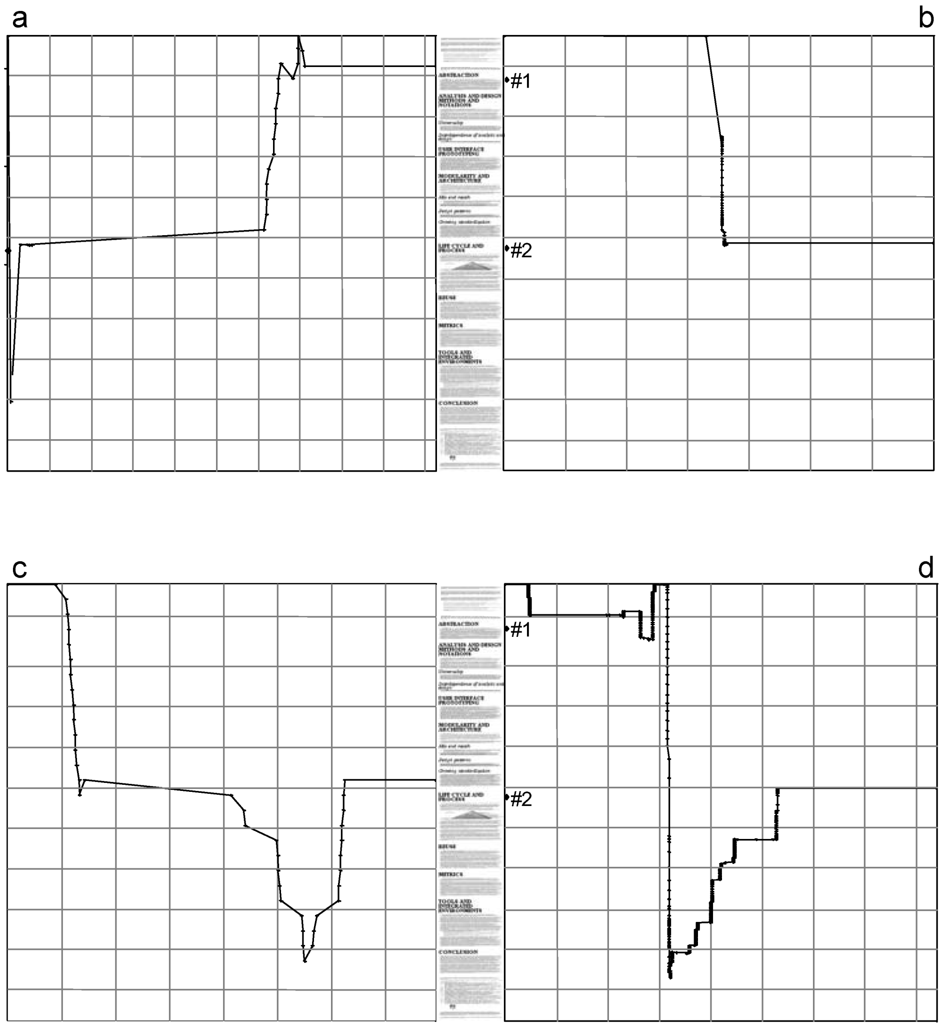


Fig. 9. Example of typical progression maps for a specific question in a question-answering task. Panel **a** shows a subject who locates both targets, using a direct jump to find the first target. Panel **b** shows a subject who locates only one target, after having had visible the first part of the document for half the task completion time. Panel **c** shows a typical reading pattern, where target #2 is located, the document is further explored, and then the subject returns to the target found. Panel **d** shows a subject who relatively quickly makes first contact with target #1, explores the document further and then locates target #2. The task completion times are **a** 3.0 min., **b** 4.0 min., **c** 4.6 min., and **d** 4.8 min.

accounts for 30% of the returns. The use of the overview pane for returning to targets suggests that subjects remember the position of previously visited parts of the document on the overview pane.

4.2.2 *Further Explorations.* When subjects, after having made first contact, stop target reading and navigate to a non-target area, we say they make *further explorations*. We do not consider it a further exploration if the subject

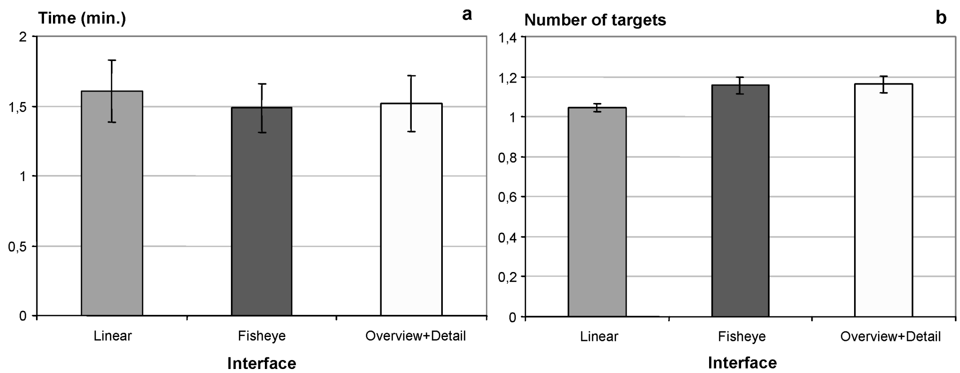


Fig. 10. Time to first contact and number of targets reached. Panel a shows the average time to first contact in the three interfaces. Panel b shows the average number of targets found in the three interfaces. Error bars show the standard error of the mean. This figure only includes tasks in which one or more targets are reached (N = 335).

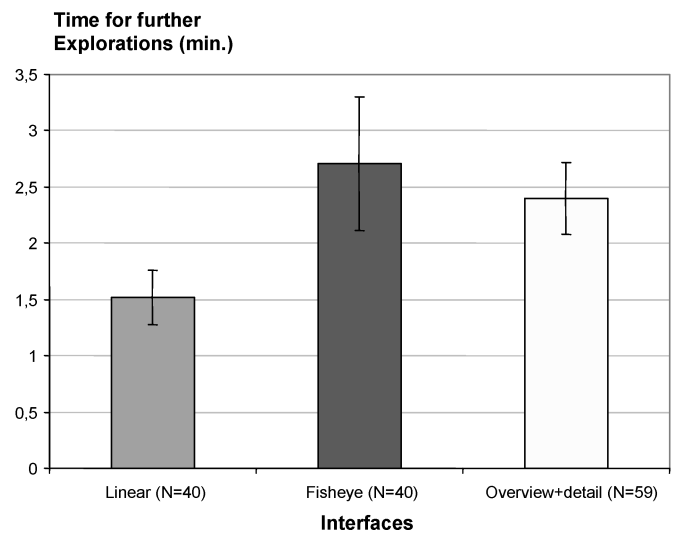


Fig. 11. Further explorations. This figure shows the frequency of tasks with one or more further explorations (shown in parenthesis under the bars) and the average duration of further explorations in the three interfaces. Error bars show the standard error of the mean.

navigated directly to another target. Also, brief navigation to another part of the document is not considered a further exploration. Operationally, navigation to a non-target area is only considered a further exploration if it lasts more than 10 seconds.

Figure 11 shows the average number of question-answering tasks in which subjects explore the document further. The number of further explorations were significantly different between interfaces, $\chi^2[2, N = 354] = 7.59$, $p < .05$. Subjects explored the document further in 48% more tasks in the overview+detail interface than in either the linear or the fisheye interface. Figure 11 also shows that subjects explore the document for different lengths of time,

$F[2, 313] = 3.87$, $p < .05$. Compared to the linear interface ($M = 1.6$ min., $SD = 1.56$), significantly more time is used exploring the document in the overview+detail interface ($M = 2.4$ min, $SD = 2.46$, $F[1, 313] = 7.46$, $p < .01$). There is no significant difference between the linear interface and the fisheye interface ($M = 2.7$ min, $SD = 3.75$, $F[1, 313] = 3.26$, $p > .05$). In summary, further exploration is more frequent and lasts longer with the overview+detail than with the linear interface.

The observation that a greater number of further explorations is made with the overview+detail interface than with the linear interface is also supported by Figure 10, which shows that the number of targets found differed between interfaces, $F[2, 313] = 6.97$, $p < .001$. A linear contrast between interfaces suggests that 10% more targets were located in the overview+detail interface ($M = 1.16$, $SD = .44$) than in the linear interface ($M = 1.05$, $SD = .21$), $F[1, 313] = 13.2$, $p < .001$. Note that only one third of the tasks contains multiple targets. The difference in number of targets located suggests that subjects keep exploring the document in the overview+detail condition, even when a satisfactory answer has been found.

One final point concerns the techniques subjects use for initiating further explorations. In the overview+detail interface, subjects often clicked on the overview pane to navigate to the area clicked on. The progression maps show that subjects use jumping on the overview to begin further exploration 22 times (26%). That number is similar to the number of times further exploration begins by subjects scrolling up in the text (also 26%) and less than the number of times subjects scroll down (38%).

5. USABILITY MEASURES

This section presents the differences in measures of usability between interfaces. We use the reading patterns presented in the previous section to explain these differences.

5.1 Grades and Incidental Learning

Figure 12 shows the average grade and incidental-learning score for the three interfaces. For essay tasks, we found a significant influence of interface on the average grade obtained, $F[2, 32] = 4.16$, $p < .05$. Linear contrasts show that essay tasks solved with the overview+detail interface ($M = 2.47$, $SD = .84$) on average got half a grade higher compared to the linear ($M = 2.00$, $SD = .86$, $F[1, 32] = 5.26$, $p < .05$) and the fisheye interface ($M = 1.95$, $SD = .78$, $F[1, 32] = 7.10$, $p < .05$). Based on the reading patterns, we have no direct explanation for this finding. However, the question-answering tasks suggest that subjects are able to use the overview pane to navigate and that they remember the position of information on the pane.

The number of correctly answered incidental learning questions differed significantly between interfaces, $F[2, 32] = 6.80$, $p < .01$. Subjects correctly answered fewer questions in the fisheye interface ($M = 3.42$, $SD = 1.22$) compared to the linear interface ($M = 4.20$, $SD = 1.24$, $F[1, 32] = 8.22$, $p < .01$) and the overview+detail interface ($M = 4.58$, $SD = 1.22$, $F[1, 32] = 11.83$, $p < 0.01$). On

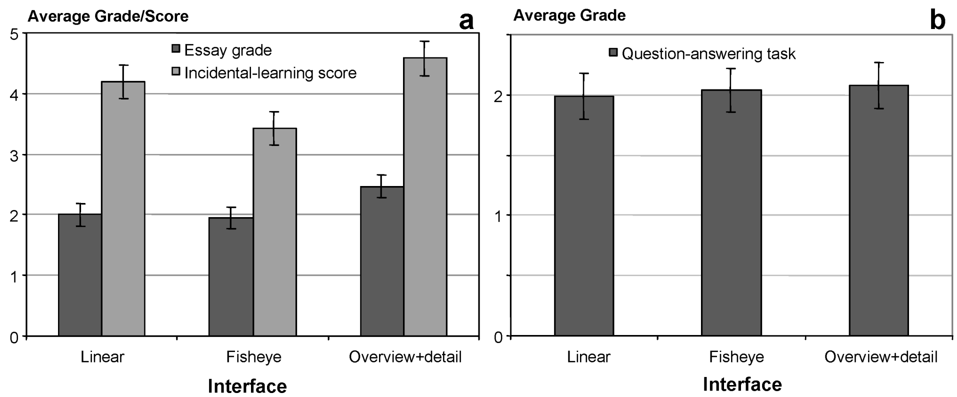


Fig. 12. Grades and incidental learning. Panel a shows for the essays tasks the average incidental learning score and grade for the three interfaces (N = 58). Panel b shows for the questions-answering tasks the average grade for each question (N = 354). Error bars show the standard error of the mean.

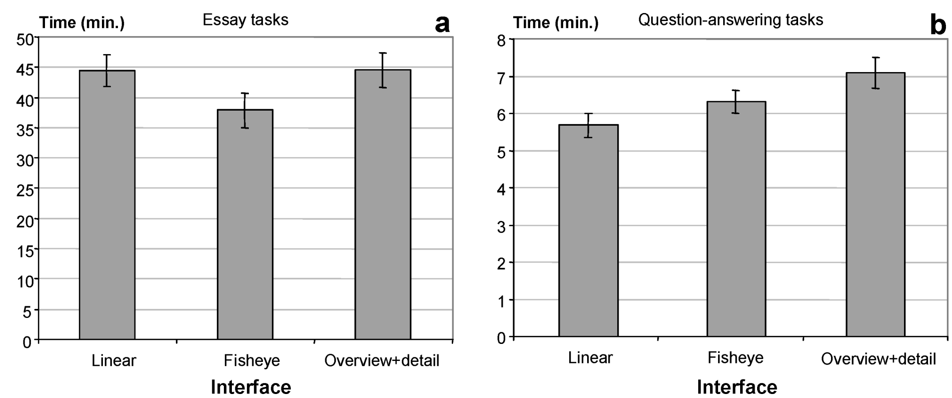


Fig. 13. Task completion time. Panel a shows the average task completion time for the essay task for the three interfaces. Panel b shows the average task completion time for question-answering tasks. Error bars show the standard error of the mean.

the average, around one question fewer was correctly answered when subjects used the fisheye interface. The visibility maps suggest that subjects pay less attention to initially collapsed sections, thereby missing information for some incidental learning questions. The overview-oriented reading style of the fisheye interface apparent from the analysis of reading modes suggests a similar reason for subjects' low incidental-learning score.

For question-answering tasks, we found no difference between interfaces for the grades given to the tasks, $F[2, 313] = .18$, $p > .5$.

5.2 Task Completion Time

Figure 13 shows the task completion time for essay and question-answering tasks. For essay tasks, we find a difference in task completion time, $F[2, 32] = 4.92$, $p < .05$. A linear contrast analysis shows that the fisheye interface

($M = 37.4$ min., $SD = 12.4$) was approximately 16% faster than the linear interface ($M = 44.4$ min., $SD = 11.9$, $F[1, 32] = 8.13$, $p < .01$) and the overview+detail interface ($M = 44.5$ min., $SD = 12.2$, $F[1, 32] = 6.51$, $p < .05$). The reason why subjects are faster with the fisheye interface appears to be the overview-oriented reading style and the short time subjects look at initially collapsed sections. Note that subjects using the fisheye interface have to expand most of the sections in the document—what is different from the other interfaces is the length of time these sections are visible.

For the question-answering tasks we also found a significant difference between interfaces, $F[2, 313] = 4.235$, $p < .05$. The overview+detail interface ($M = 7.1$ min., $SD = 4.1$) was 20% slower then the linear interface ($M = 5.9$ min., $SD = 3.5$, $F[1, 313] = 8.33$, $p < .01$). Note that, as explained in Section 3.2, we imposed no time limit on the subjects' work with the tasks. According to the reading patterns, this time difference is not due to difficulty in locating a target. However, in the reading patterns of the overview+detail interface further explorations are more frequent and last longer. Also, the number of targets found with the overview+detail interface is higher than in the other interfaces. Subjects initiated many further explorations by directly clicking on the overview pane.

5.3 Satisfaction

Nineteen of the subjects preferred using the overview+detail interface; one subject preferred the linear interface. In their motivation for preferring the overview+detail interface, 10 subjects mentioned the overview of the documents structure and headings as an important reason; six subjects mentioned that the overview+detail interface supports easy navigation; and five subjects liked that highlighted words show up in the overview pane. Fourteen subjects mentioned that they found it hard to overview the document using the linear interface. With respect to the fisheye interface, nine subjects commented that they did not like to depend on an algorithm to determine which parts of the document should be readable. Subjects were divided as to whether the fisheye made it easier ($N = 5$) or harder ($N = 2$) to get an overview of a document. Figure 14 shows the subjects' answers to the questionnaires received after using each of the interfaces. We compared interfaces using paired t-tests with a Bonferroni-adjustment of $0.05/12^{*}3 \approx .0013$. The overview+detail interface scored higher than the two other interfaces on questions about overall satisfaction, and on the dimensions terrible–wonderful and frustrating–pleasant. Subjects scored the fisheye interface lower compared to the overview+detail interface on the dimension confusing–clear. Subjects also scored the overview+detail interface higher than the linear interface on the question of whether the documents were easy or hard to overview. We found no difference for the questions intended to investigate whether the subjects' perception of their tasks differed between interfaces (question 7 and 8 in Figure 14).

Subjects' satisfaction with the overview is supported by several of the reading patterns. The overview pane supports jumping directly to targets; it helps returning to previously visited parts of the document; and it invites and supports

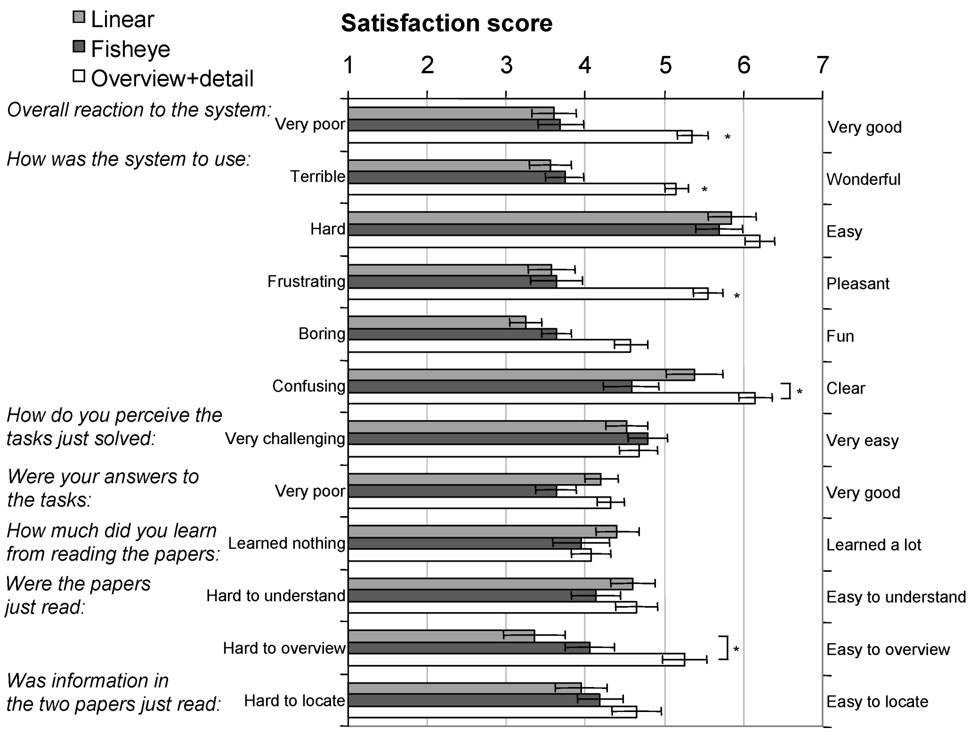


Fig. 14. Satisfaction. To the left, in italics, is shown the questions subjects were given. Negative and positive concepts on a seven-point semantic differential are shown to the left and right of the chart. Low scores were given to the negative concept of the differential scale. The bars show the average satisfaction scores for the three interfaces. An asterisk (*) denotes a significant difference using a Bonferroni adjustment of .0013.

further explorations. Subjects using the fisheye interface depend extensively on the algorithm that determines which sections to collapse initially, even though subjects do not trust this algorithm.

6. DISCUSSION

The overview+detail interface was slow for question-answering tasks. This result contradicts the expectations raised in previous work (see Section 2) and our hypothesis (see Section 3) about the usability of overview+detail interfaces. But our result is similar to the empirical results of Dee-Lucas and Larkin [1995] and Hornbæk et al. [2002] who also found that overviews may lead to higher task completion times. We find that the reading patterns offer an explanation of this surprising result. First of all, reading patterns show that, across interfaces, subjects make first contact after comparable times. Thus, whatever slows down subjects happens after the first contact with a target. The reading patterns also suggest that further explorations were more frequent in the overview+detail interface than in the two other interfaces; the explorations also took more time than in the linear interface. Thus, the frequency and duration of further explorations seems to be the main reason for the longer task completion time.

But why does the overview+detail interface invite subjects to make further explorations? Here, it seems noteworthy that further explorations were often initiated by clicking on the overview pane. These observations support the explanation previously proposed [Hornbæk and Frøkjær 2001] that the overview pane grabs subjects' attention, and thereby leads them to explorations that strictly speaking are unnecessary. Seemingly, further explorations happen because of the visual appearance of the overview and because of the navigation possibilities afforded by the ability to click the overview pane. More broadly, the influence from the overview pane upon reading activity also suggests that reading is more unpredictable and shaped by situation-dependent inspiration and associations. For example from an interface, as opposed to the description offered by Guthrie's [1998] model of locating information in documents, which suggests a rational, goal-oriented process. Perhaps the lack of time limit in our study plays a role in this observation.

In contrast to the time spent, perhaps unnecessarily, on further explorations, the overview+detail interface led to higher quality essays, and subjects strongly preferred this interface. It is not clear from the reading patterns why subjects write better essays. One explanation might be that the overview pane may indirectly have helped subjects to organize and recall text. The subjective satisfaction data suggest that the overview+detail interface supported navigation, was easier to overview, invited explorations, seemed clear and convenient to use, and supported jumping directly to previously read text. The data suggest that subjects are free to concentrate on reading instead of operating the interface. The higher subjective satisfaction might also, through higher motivation, affect the grades given to essays. Thus, although the overview+detail interface might be slower for question-answering tasks, we think designers would be well advised to use overview+detail interfaces for electronic documents.

For essay tasks, the fisheye interface was approximately 16% faster than the other interfaces. Reading patterns suggest two reasons for this. First, subjects used more time in the initial orientation mode and less time in the linear read-through mode. These reading patterns confirm our hypothesis that the fisheye interface shortens navigation time by supporting an overview-oriented reading style. Second, reading patterns show that subjects opened almost all collapsed sections with the fisheye interface, suggesting that the interface does not change *which* parts of the document subjects have visible. However, what does seem to change is the length of time subjects can see the different parts of the document. Subjects spent approximately 30% less time on the initially collapsed sections compared to the other interfaces. We were surprised that the initial status of sections so strongly influenced the length of time they were visible, especially as many subjects expressed a lack of trust in the algorithm. One explanation might be that subjects assume that the contents of the initially collapsed sections are not important, rather than judging the importance of sections from what they read in those sections. This behaviour is akin to premature cognitive commitment [Langer 1991], where humans commit themselves to one view on or use of information and at a later time fail to reconsider their commitment. Independently of the validity of this explanation, the fisheye interface seems to fundamentally change the way subjects perceive and interact with documents.

Usability data and reading patterns together suggest that fisheye interfaces should be used mainly for time-critical tasks. But depending on the task context, the finding that subjects obtain lower incidental learning scores with the fisheye interface might be of importance. The fisheye interface therefore seems more appropriate for tasks where a detailed understanding of the document is not the main aim, for example in relevance judgments as carried out in information retrieval systems, when users must judge whether it is worthwhile to download or thoroughly read a document.

The linear interface is in many ways inferior with respect to usability to the two other interfaces. For essay tasks, the overview+detail interface leads to higher essay grades; the fisheye interface is faster. The linear interface scores lower than the overview+detail interface on several satisfaction questions, including one concerning how easy it is to overview a document. We therefore recommend that designers rely less on this interface type and use the overview+detail and, in special cases, the fisheye interface.

Our visualizations of reading patterns suggest four interesting observations in addition to those mentioned above. First, flipping through the document might give subjects an initial indication of its length, structure and key elements. Many subjects seem to like doing a flip-through to set the scene for a more careful reading of the documents. The flip-through behaviour seems to be one strategy that subjects develop to cope with the low tangibility of electronic documents [Hansen and Haas 1988; O'Hara and Sellen 1997]. Second, in question-answering tasks the reading patterns show how subjects used the overview-pane to directly jump back to previously visited targets. Thus, the overview pane supports helps readers memorize important document positions, perhaps analogously to the way readers remember the position of information in paper documents [Rothkopf 1971]. This ability of subjects to mentally couple the overview pane with the text they have read is one of the most important aspects of the overview, and may influence essay writing. Third, our observations on reading patterns confirm and extend previous research on reading, outlined in Section 2. Non-linear navigation occurred extensively, but mostly at the beginning and end of the reading activity. We also found, similarly to Foltz [1996], that most of the reading time consisted of linear reading through the document. Our data suggests that these seemingly very different reading patterns are integrated in reading activity and may be used to differentiate different modes of the reading activity: initial orientation and review modes are characterized by non-linear navigation; the main work with the document is a linear read-through. Fourth, we found large differences between individuals in reading strategies, as did, for example, Goldman and Saul [1990]. However, in this paper we concentrated on examining the influence of interface on reading patterns, not on describing individual differences.

Concerning methods for studying reading patterns, our idea of progression maps offers an intermediate level of analysis of user behaviour in reading electronic documents, between coarse measures such as task completion time, and fine-detail analysis such as eye-tracking (e.g. as in Zellweger et al. [2000]). The most important limitation of our method is that we can only register what parts of the documents are visible, not what subjects actually looked at.

In relation to the aims set forth in the introduction to this paper, we have investigated both the usability and the reading process in visualizations of electronic documents. To follow up this investigation, we suggest four areas of further research. First, we need to improve visualizations of electronic documents. The algorithm underlying the fisheye interface may be improved based on our descriptions of reading patterns. The overview pane may also benefit from more information-rich semantic zooming. Second, we have not investigated the potential and problems in combining interfaces, either by having users choose the most appropriate interface for a certain task, or by combining the fisheye interface's support for expanding and collapsing text with the overview+detail interface's overview pane. Though a combination of interfaces seems attractive, experiments are needed to uncover whether the problems of having different ways of performing the same operations (e.g. as described by Raskin [2000]) outweigh a possible positive influence on usability. Third, we need a more thorough study of reading activity during actual work as performed by subjects who have gained full familiarity with the experimental interfaces. Fourth, we need some better theories of how attention is shaped by visualizations. The role of the overview in triggering further explorations is not well described by any theory of information visualization of which we are aware.

7. CONCLUSION

In an experiment, we compared three interfaces for electronic documents. Two of the interfaces were based on overview+detail and fisheye visualizations; the third was a linear interface that served as a baseline. Subjects in the experiment answered questions and wrote essays about scientific documents. In an attempt to better understand how the interfaces supported reading and to understand the differences in measures of usability between interfaces, we created visualizations of subjects' reading activity by two kinds of maps. Progression maps were used to depict how the reading progressed; visibility maps were used to compare the average length of time different parts of the document were visible. From these visualizations we describe how interfaces shape subjects' reading patterns.

Subjects clearly preferred the overview+detail interface, especially because of the overview gained and the ease of combining navigation with reading. When using this interface, subjects receive higher grades for their essays. For question-answering tasks, the progression maps show that subjects with the overview+detail interface explore the document more often than with the other interfaces. Consequently, subjects use more time answering questions. The visibility maps reveal that subjects with the fisheye interface have the parts of the document that are not initially readable visible for less time. With the fisheye interface, subjects also read the documents using an overview-oriented reading style. Therefore, subjects read faster with this interface, but display lower incidental learning.

As for the practical issue of supporting reading, visualizations such as the overview+detail or fisheye interface improve the usability of electronic documents. However, such visualizations also change how subjects read documents.

The most common interface in practical use, the linear interface, was inferior on most usability aspects compared to the other two interfaces. Visualizations are thus recommended to developers as usable interfaces for electronic documents. For researchers, further improvement of visualization of electronic documents is feasible, as are use of progression and visibility maps to study and improve reading activity.

ACKNOWLEDGMENTS

We thank Peter Naur and Morten Hertzum for comprehensive comments on a draft of this paper; Aran Lunzer helped with thorough revisions of our English. The anonymous reviewers also helped greatly in improving the clarity of the paper.

REFERENCES

ANDERSON, T. H. AND ARMBRUSTER, B. B. 1982. Reader and text-studying strategies. In *Reading Expository Material.* W. OTTO AND S. WHITE, Eds. New York, NY: Academic Press. 219–242.

BAZERMAN, C. 1988. *Shaping Written Knowledge.* Madison, WI: The University of Wisconsin Press.

BISHOP, A. P. 1995. Working towards an understanding of digital library use. *D-Lib Magazine*, October 1995, http://www.dlib.org/dlib/october95/10bishop.html.

BISHOP, A. P. 1999. Document structure and digital libraries: how researchers mobilize information in journal articles. *Info. Proc. Manag. 35*, 255–279.

BOGURAEV, B., KENNEDY, C., BELLAMY, R., BRAWER, S., AND WONG, Y. Y. 1998. Dynamic presentation of document content for rapid on-line skimming. In *AAAI Spring 1998 Symposium on Intelligent Text Summarization,* E. HOVY AND D. RADEV, Eds., AAAI Press, Stanford CA, 109–118.

BRADOW, R., MITZE, K., AND RAU, L. F. 1995. Automatic condensation of electronic publications by sentence selection. *Info. Proc. Manag. 31*, 5, 675–685.

BYRD, D. 1999. A scrollbar-based visualization for document navigation. In *Proceedings of Digital libraries '99,* Berkeley CA, Aug. 1999, N. ROWE AND E. A. FOX, Eds., ACM Press, New York, NY, 122–129.

BYRNE, M. D., JOHN, B. E., WEHRLE, N. S., AND CROW, D. C. 1999. The tangled web we wove: a taskonomy of WWW use. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Pittsburgh PA, May 1999, M. G. WILLIAMS AND M. W. ALTOM, Eds., ACM Press, New York, NY, 544–551.

CHEN, C., AND RADA, R. 1996. Interacting with hypertext: A meta-analysis of experimental studies. *Hum.-Comput. Interact. 11*, 2, 125–156.

COHEN, J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

DEE-LUCAS, D. AND LARKIN, J. H. 1995. Learning from electronic texts: Effects of interactive overviews for information access. *Cogn. Instruct. 13*, 3, 431–468.

DILLON, A. 1994. *Designing Usable Electronic Text.* London: Taylor & Francis.

DYSON, M. C. AND HASELGROVE, M. 2001. The influence of reading speed and line length on the effectiveness of reading from screen. *Int. J. Hum.-Comput. Studies*, 4, 585–612.

EGAN, D. E., REMDE, J. R., GOMEZ, L. M., LANDAUER, T. K., EBERHARDT, J., AND LOCHBAUM, C. C. 1989. Formative design-evaluation of SuperBook. *ACM Trans. Info. Syst. 7,* 1, 30–57.

EICK, S. G., STEFFEN, J. L., AND SUMNER, E. E. 1992. Seesoft–A tool for visualizing line oriented software statistics. *IEEE Trans. Soft. Eng. 18,* 11, 957–968.

FOLTZ, P. W. 1996. Comprehension, coherence and strategies in hypertext and linear text. In *Hypertext and Cognition.* ROUET, J. F., LEVONEN, J. J., DILLON, ANDREW P., AND SPIRO, R. J., Eds. New Jersey: Lawrence Erlbaum Associates. 109–136.

FURNAS, G. W. 1986. Generalized fisheye views. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, Boston MA, Apr. 1986, M. MANTEI AND P. ORBETON, Eds., ACM Press, New York, NY, 16–23.

GINSBURG, A., MARKS, J., AND SHIEBER, S. 1996. A viewer for postscript documents. In *ACM Symposium on User Interface Software and Technology*, Seattle, WA, Nov. 1996, D. KURLANDER, M. BROWN AND R. RAO, Eds., ACM Press, New York, NY, 31–32.

GOLDMAN, S. R. AND SAUL, E. U. 1990. Flexibility in text processing. *Learning and Individual Differences, 2*, 2, 181–219.

GRAHAM, J. 1999. The reader's helper: a personalized document reading environment. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Pittsburgh PA, May 1999, M. G. WILLIAMS AND M. W. ALTOM, Eds., ACM Press, New York, NY, 481–488.

GUTHRIE, J. 1998. Locating information in documents: examination of a cognitive model. *Reading Research Quarterly, 23*, 178–199.

HANSEN, W. J. AND HAAS, C. 1988. Reading and writing with computers: a framework for explaining differences in performance. *Comm. ACM, 31*, 9, 1080–1089.

HERTZUM, M. AND FRØKJÆR, E. 1996. Browsing and querying in online documentation: a study of user interfaces and the interaction process. *ACM Trans. Comput.-Hum. Interact. 3*, 2, 136–161.

HERTZUM, M., LALMAS, M., AND FRØKJÆR, E. 2001. How are searching and reading intertwined during retrieval from hierarchically structured documents? In *Proceedings of the IFIP TC 13 International Conference on Human-Computer Interaction*, Tokyo, Japan, Jul. 2001, M. KUROSU, Ed., IOS Press, Amsterdam, 537–544.

HOLMQUIST, L. E. 1997. Focus+context visualization with flip zooming and the zoom browser. In *CHI '97 Extended Abstracts*, Atlanta GA, Mar. 1997, S. PEMBERTON, Ed., ACM Press, New York, NY, 263–264.

HORNBÆK, K., BEDERSON, B., AND PLAISANT, C. 2002. Navigation patterns and usability of zoomable user interfaces with and without an overview. *ACM Trans. Comput.-Hum. Interact. 9,* 4, 362–389.

HORNBÆK, K. AND FRØKJÆR, E. 2001. Reading electronic documents: The usability of linear, fisheye, and overview+detail interfaces. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI Letters *3*, 1, 293–300.

HORNEY, M. AND ANDERSON-INMAN, L. 1994. The ElectroText project: Hypertext reading patterns of middle school students. *J. Educ. Multimedia Hypermedia, 3*, 1, 71–91.

KAUGARS, K. 1998. Integrated multi scale text retrieval visualization. In *CHI '98 Summary*, Los Angeles CA, Apr. 1998, C.-M. KARAT AND A. LUND, Eds., ACM Press, New York, NY, 307–308.

KUPIEC, J., PEDERSEN, J., AND CHEN, F. 1995. A trainable document summarizer. In *Proceedings of the Annual ACM Conference on Research and Development in Information Retrieval*, Seattle WA, Jul. 1995, E. A. FOX, P. INGWERSEN AND R. FIDEL, Eds., ACM Press, New York, NY, 68–73.

LANGER, E. J. 1991. *Mindfulness.* London: Harvill.

MARCUS, A. 1992. *Graphic Design for Electronic Documents and User Interfaces.* New York, NY: ACM Press.

MILLS, C. B. AND WELDON, L. J. 1987. Reading text from computer screens. *ACM Comput. Surv. 19*, 4, 329–357.

MUTER, P. 1996. Interface design and optimization of reading of continuous text. In *Cognitive Aspects of Electronic Text Processing.* H. VAN OOSTENDORP AND S. DE MUL, Eds. Norwood, NJ: Ablex.

O'HARA, K. AND SELLEN, A. 1997. A comparison of reading paper and on-line documents. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, Atlanta, GA, Mar. 1997, S. PEMBERTON, Ed., ACM Press, New York, NY, 335–342.

PÁEZ, L. B., DA SILVA-FH., J. B., AND MARCHIONINI, G. 1996. Disorientation in electronic environments: A study of hypertext and continuous zooming interfaces. In *Proceedings of the 59th Annual Meeting of the American Society for Information Science,* Baltimore, MD, Oct. 1996, HARDING, S., Ed., 58–66.

PLAISANT, C., CARR, D., AND SHNEIDERMAN, B. 1995. Image browsers: Taxonomy, guidelines, and informal specifications. *IEEE Soft. 12*, 2, 21–32.

RASKIN, J. 2000. *The Humane Interface*. Reading MA: Addison-Wesley.

ROBERTSON, G. G. AND MACKINLAY, J. D. 1993. The document lens. In *Proceedings of the ACM Symposium on User Interface Software and Technology,* Atlanta GA, Nov. 1993, S. HUDSON, R. PAUSCH, B. V. ZANDEN AND J. FOLEY, Eds., ACM Press, New York, NY, 101–108.

ROSENTHAL, R., ROSNOW, R. L., AND RUBIN, D. B. 2000. *Contrasts and Effect Sizes in Behavioral Research.* Cambridge: Cambridge University Press.

ROTHKOPF, E. 1971. Incidental memory for location of information in text. *J. Verbal Learn. Verbal Behav. 10,* 608–613.

SCHRIVER, K. 1997. *Dynamics in Document Presentation.* New York/London: John Wiley & Sons.

SELLEN, A. AND HARPER, R. 1997. Paper as an analytic resource for the design of new technologies. In *Proceedings of the ACM Conference on Human Factors in Computing Systems,* Atlanta GA, Mar. 1997, S. PEMBERTON, Ed., ACM Press, New York, NY, 319–326.

SELLEN, A. AND HARPER, R. 2002. *The Myth of the Paperless Office.* Cambridge, MA: MIT Press.

WRIGHT, P. 1987. Reading and writing for electronic journals. In *Executive Control Processes in Reading.* B. K. BRITTON AND S. M. GLYNN, Eds.Hillsdale, NJ: Lawrence Erlbaum Associates. 23–55.

ZELLWEGER, P. T., REGLI, S. H., MACKINLAY, J. D., AND CHANG, B.-W. 2000. The impact of Fluid Documents on reading and browsing: An observational study. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, the Hague, the Netherlands, Apr. 2000, T. TURNER AND G. SZWILLUS, Eds., ACM Press, New York, NY, 249–256.

Received November 2001; revised July 2002; accepted January 2003