John Benjamins Publishing Company

Information Design Journal 20(1), **2–15**

DOI: 10.1075/idj.20.1.01far

David K. Farkas and Christopher A. Raleigh

Designing documents for selective reading

> "Some books are to be tasted, others to be swallowed, and some few to be chewed and digested: that is, some books are to be read only in parts, others to be read, but not curiously, and some few to be read wholly, and with diligence and attention."
>
> Sir Francis Bacon, *Of Studies*, 1625

Keywords: literacy, information design, hypermedia, summarization, selective reading

There are strong indications that many people increasingly resist reading medium-to-long documents. It is therefore important to contribute to the long-term viability of longer documents by providing better support for selective reading. Readers may be more willing to read longer documents knowing they have ample and near-seamless choices regarding which topics they can read and the level of detail at which they can read a particular topic. To design for selective reading requires an understanding of how readers deal with incomplete information and the concepts of prerequisite information and dependency relationships. Three broad approaches can be identified: building supported reading pathways, modularization, and summarization.

All of us regularly engage in selective reading within documents, and our documents offer a wide range of familiar affordances to support selective reading. Traditional documents, such as reports and books, provide support for selective reading through such features as explanatory footnotes, appendices, and optional chapters. In the digital media, a very large class of documents including websites and help systems are strongly nonlinear in their construction and invite readers to descend any branch of a wide hierarchy. Now may be the time, however, to focus our attention on selective reading as a design problem and to devise a broader range of ways to support it.

Why now? There are strong indications that many people are engaging less frequently in the sustained linear reading of medium-to-long documents (Carr, 2010; Liu, 2005; Palfrey & Gasser, 2008; Nicholas et al., 2008; Menzies & Newson, 2007; Richtel, 2010). And why should this be true? First, we appear to be living in an era of information overload (Eppler & Mengis, 2004), where many people feel there is more information they must or want to read (or otherwise process) than they have time for (Thomas et al., 2006; Menzies & Newson, 2007; Wurman, 2001; Schick, Gordon, & Haka, 1990). Second, many people, especially digital natives, prefer to read large numbers of short documents such as text messages and social media posts rather than longer documents (Palfrey & Gasser, 2008; Richtel, 2010).

One response to this likely change in reading habits is a shift to the publication of shorter documents, as can be seen in such new publishing programs as Byliner (Byliner, 2012), Amazon Singles (Dignan, 2010), and SpringerBriefs (Springer, n.d.). But assuming, as we do, that in many situations the world still needs longer documents in order to communicate thoroughly and meaningfully about complex topics, better support for selective reading may contribute to the long-term viability of longer documents. Readers may be more likely to read longer documents if they know they have ample and near-seamless choices regarding which topics within the document they can read and the level of detail at which they can read a particular topic.

Our goal here is to enrich the present understanding of how to design for the selective reading of medium-to-long non-fiction documents belonging to a wide range of genres that inform and educate about a complex body of ideas. We briefly explain how readers deal with incomplete information and how authors manage dependency relationships among the ideas they present. Then we provide a wide-ranging survey of how documents can be designed for selective reading. This survey substantially updates a survey with similar aims written in 1988 (Holland, Charrow, & Wright). We identify and describe three broad approaches to designing for selective reading, and we provide examples of each. These approaches are (1) building supported reading pathways, (2) modularizing content, and (3) summarization.

The rhetoric of dependency relationships

To design for selective reading, there are certain things we must know about how human beings read and how documents must be structured to enable efficient information processing.

How we read

The reading process consists both of taking in new information and applying existing knowledge (Kintsch, 1998; LaBerge & Samuels, 1974). As we absorb new information, we incorporate it into long-term memory and integrate it into our existing knowledge in the form of schemata (Anderson, 2004).

The reading process is robust: we are good at making inferences. That is, if there is a gap in our existing schema, we will draw upon the most relevant information we possess (McNamara & Kintsch, 1996). If we find that an item of relevant information is missing from what we are reading, we can often hold the larger idea in memory and wait for the missing piece. For example, someone who does not know soccer but encounters a routine newspaper article on a particular soccer game will probably be able to understand the gist of the article, especially if that person (as most of us do) has some familiarity with hockey, basketball, or another sport that involves scoring goals or baskets. If the article mentions corner kicks, this reader is apt to figure out that a corner kick is a good scoring opportunity. Furthermore, it is very possible that the detailed account of a particular corner kick will help our reader better understand corner kicks in general. Of course, the robust nature of the reading process has limits. If a lay reader picks up an academic journal article dealing with theoretical physics, the much larger deficit in background knowledge will quickly cause disruption and breakdown in the reading process.

Managing dependency relationships

The reading process greatly influences how we, as writers, sequence the ideas we present in discourse. To minimize the burden of reading and to create a satisfying reading experience, we carefully organize documents to

provide a smooth progression from the initial knowledge level we expect our readers to possess to the full set of ideas we mean to convey (Chambliss & Calfee, 1989).

Let's first consider book-length documents such as complex expositions and arguments that are strictly linear in their construction – that is, designed to be read from beginning to end. All the chapters (and, within a chapter, all the sections) have a "building block" relationship. Each is a prerequisite to the next as the author carefully manages the expanding information context (what the reader is presumed to have learned from the text). Authors strive to avoid gaps in the presentation, in particular missing or delayed information. Authors also manage the reading experience in many other ways, such as providing previews to activate the relevant schema that readers may have and providing headings to show the hierarchical relationship among the major ideas and to facilitate scanning for information (Lorch, 1989).

While the author of a linear document must manage the evolving context of information carefully, the reader takes on a complementary responsibility to read linearly. If the reader violates their side of this implied contract by skipping sections or chapters, the author's responsibility ends or is greatly reduced. Few people would sympathize with a reader who jumps around within a complex exposition or argument and then complains that the book is badly written and doesn't make sense. In much the same way, a hiker in the heavily visited parts of a national forest legitimately expects the trails to be reasonably easy to follow. But if the hiker leaves the trail to bushwhack through the forest and then gets into trouble, the forest managers bear little responsibility.

Non-linear documents offer one or more optional pathways, and each pathway creates a new context and, hence, more complex dependency relationships. An author, for example, might indicate that Chapter 4 is an optional chapter, perhaps a more detailed look at the topic explained in Chapter 3. The book now offers two supported pathways: two contexts have opened for the author to monitor and manage. One context consists of the ideas in Chapters 1–3: the other consists of the ideas in Chapters 1–4. Monitoring and managing these two contexts, while necessary, is not difficult: the author must simply ensure that nothing in Chapter 5 or beyond requires an idea explained only in Chapter 4 for a smooth and productive reading experience.

In genres such as textbooks and computer-based tutorials, multiple contexts (variations in what the reader is presumed to have learned from the text) often arise. Monitoring these multiple contexts can be an arduous or even impossible task for the author. Fortunately, however, because the reading process is robust, multiple contexts can often be managed without monitoring them individually. First, some domains are built in large part upon everyday knowledge that most readers can supply. But even in highly specialized domains (in which dependencies are inherently strong) an author can offer ample background explanations in lieu of monitoring each individual context.

As an analogy, consider a college instructor who is teaching Biology 103 to students who have taken one of three sections of Biology 101 and one of two sections of Biology 102. These five sections have all been taught by different instructors who have covered somewhat different content. The Biology 103 instructor is therefore responsible for six different contexts (3 x 2). Rather than monitoring these specific contexts (maintaining an ongoing awareness of the individual information gaps of the six subsets of students who have studied in the different sections), the instructor manages the dependency relationships in a generalized "shotgun" style. That is, whenever she introduces a new topic, she will provide ample background information so as to be reasonably confident that her students will understand the new topic regardless of which section of Biology 101 and 102 each student has taken.

Modularization is a very different means of managing dependency relationships. Here discrete chunks of content are written as free-standing modules for which there are no prerequisites (other than a certain baseline level of general knowledge and reading ability). The reader can choose which modules to read or skip without any loss of context (that is, without reaching a place in the document where prerequisite information is missing). For example, Farkas recalls that his high school biology textbook took the form of ten modular chapters, each dealing with a system of the human body (e.g., the digestive system). Because of the modular design, the chapters could be read in any order (except that the chapter on the reproductive system was printed in a separate booklet the class never saw).

Dependency relationships are rhetorical

Dependency relationships are not an inherent characteristic of texts but rather are rhetorical, dependent on the background knowledge, needs, and temperament of individual readers. If a mathematics instructor peruses a newly published algebra textbook, let us say a textbook in which the chapters are designed for linear reading, the instructor, because of her advanced knowledge of the domain of algebra, can skip chapters with little or no disruption. For her – though not for her students – there are no prerequisites.

A middle school student who needs only basic information about the Normandy Invasion might stumble upon a complex, sophisticated book about World War II. The book's chapter on the Normandy Invasion includes many themes developed earlier in the book, themes that the student, looking only at this chapter, cannot possibly understand. Still, for this student's limited information needs, the chapter may prove fully adequate.

With regard to temperament, some readers have more tolerance than others for processing information for which their background is not ideal (Lee, 2005). By way of analogy, a confident student who did well in his high school biology course might register for Biology 102 and figure he can cope with whatever information he missed by not taking Biology 101.

Three approaches to designing for selective reading

We now turn to the main part of this study, where we identify and explain the three approaches to designing for selective reading: (1) building supported reading pathways, (2) modularizing content, and (3) summarization. We recognize that among the vast number of individual designs for supporting selective reading there are hybrids that span the approaches we identify. But our classification and the examples we provide should be helpful in any analysis of the many designs we do not cover and in future design work. Note that we exclude from consideration the extensive work that has been done on XML and content management systems (Albers, 2003; Boiko, 2005). We do so because content management systems are intended to create a multitude of customized documents rather than documents with

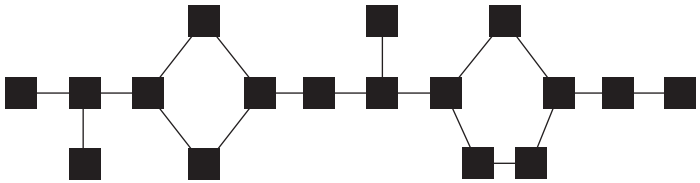


Figure 1. A node-link diagram of a multipath document with two optional nodes and two sets of alternative nodes.

special affordances for supporting selective reading. This is true even for content management systems in which users are empowered to create their own custom documents (Rockley, 2001; Severson, 2009).

Building supported reading pathways

Building supported reading pathways is sometimes described as the *layering* of information (Holland, Charrow, & Wright, 1988; Graham, 2009). Four key ways to do so are described below. These are (1) employing optional components, (2) employing alternative components, (3) multipath design, and (4) employing gateway components.

Optional components

Optional components, as noted above, include explanatory footnotes, appendices, and optional chapters. These and other optional components share an implied contract with the reader. Nothing they contain can be a prerequisite for reading the core document. Note that an optional component – for example an optional chapter or appendix – can either open up a new topic or elaborate on a topic covered in the core document.

Alternative (split-join) components

Another technique is to provide two (or more) alternative pathways. That is, the main reading pathway splits into branches, each of which contains the prerequisite information that will be necessary when the branches re-join. Consider, for example, a statistics textbook or an on-screen tutorial that offers two alternative units on multivariate regression, one with examples drawn from the social sciences and one with examples from the natural sciences. The reader is invited to read the most relevant unit knowing that either choice provides the necessary background for all subsequent units.

Multipath designs

The conjoint use of optional and alternative components, illustrated diagrammatically in Figure 1, can be usefully regarded as a distinct technique.

Textbooks, both print and electronic, are often multipath designs. Publishers and textbook authors seek broad adoption by instructors who have different preferences regarding course topics, whose students have different backgrounds, and whose school terms are of different lengths. To provide flexibility in the use of the textbook, there are often alternative (split-join) chapters and, even more frequently, optional chapters. These options are typically described in the preface or in a teacher's guide. Many computer games are also multipath designs.

Gateway components

Most computer manuals and almost all websites are structured as a hierarchy with multiple branches that readers are invited to descend. The top node of these documents – the home page of the website or the introductory chapter of the manual – functions as a gateway component. The top node provides the foundational information that enables the reader to descend productively any branch of the hierarchy.

Because few people wish to read computer manuals from cover to cover, the introductory (or gateway) chapter explains basic concepts and provides an overview of the interface. The author's intention is that the gateway chapter will prepare the reader to successfully read any of the subsequent chapters. Readers, of course, may skip even the gateway chapter and perhaps endure some disruption due to lost context when they try to use the manual. At times the gateway chapter is designed to prepare the reader to read most – but not all – of the subsequent chapters. For example, while the gateway chapter prepares the reader for Chapter 6, *Basic Styles*, both the gateway chapter and Chapter 6 are required

for Chapter 7, *Advanced Styles*. Ideally the titles of the chapters will reveal the special dependency relationship, as they do in this example.

Authors of computer manuals (and other kinds of technical documents) are especially likely to think about dependency relationships as they plan and write their manuals. Recognizing that readers favor modular designs, they will design a chapter to be modular if they can do so. Like textbooks, manuals often explain dependency relationships in a preface. Indeed, there are manuals in which complex reading pathways are revealed in a flowchart-like diagram.

Much like computer manuals, the home page of every website provides enough foundational information for the reader to navigate confidently to any second-level page or node. Also, with the benefit of this gateway information, the reader can navigate from any second-level node to any other second-level node, usually following lateral links, as shown in Figure 2 (Farkas & Farkas, 2002). These lateral links take the form of a navigation bar or similar global navigation element. Somewhat like the home page, each of the second-level nodes functions as a gateway for its own child nodes and enables readers to navigate confidently among these siblings.

An interesting and problematic issue arises with the associative links that jump across the main branches of the hierarchy at Level 3 and below, also shown in Figure 2. These associative links do not connect nodes that share a parent, and so the reader who follows an associative link is moving from one information context to another. What is the writer's contract with the reader if the writer/web designer builds such associative links, as they regularly do? The reader, we believe, has the right to expect the content of the destination page to be relevant to the page with the link. (We've all been annoyed by stupid, machine-generated links that take us to irrelevant destinations.) But there is no implied promise that the reader will be able to fully understand the content of a destination page, for that page belongs to a different context and may have different prerequisite information. In much the same way, there is no implied promise to someone who drops into an interior page of a website using a search engine. In both cases, however, the reader can expect to find the prerequisite information by navigating upward on the branch on which the destination node resides.

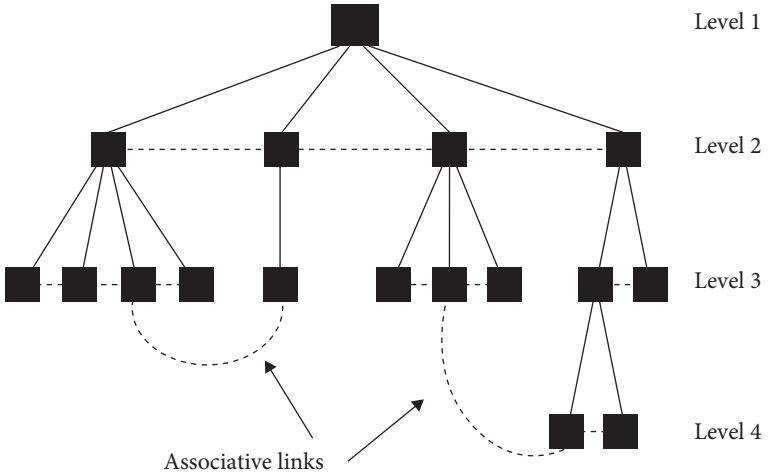


Figure 2. A node-link diagram of a website showing the primary links that comprise the website's hierarchical structure, lateral cross links that connect sibling nodes, and two associative links.

Modular design

Modular design is the second approach to designing for selective reading. Here the strategy shifts from planning specific pathways to creating content that supports all pathways and navigation choices. The paradigmatic example of modular design is the traditional encyclopedia article. No article will refer you to another article or an external information source as a prerequisite, although they may provide cross references to articles of related interest. Among the many books that consist of free-standing modules are those that follow the model exemplified by these (imaginary) titles: *Forty Ways to Save Our Planet* or *Ten Ways to Survive the Coming Economic Collapse*. The reader of these and similar books does not need to read linearly through the text. She can pick and choose among the modules.

The problem of redundancy

A major drawback of modular design is the redundancy that occurs when foundational concepts need to be repeated in more than one module. For example, assuming that *Forty Ways to Save Our Planet* includes five modules dealing with greenhouse gases, the concept of greenhouse gases must be explained, at least briefly, in each of the five modules to avoid puzzling readers who don't know this concept. Explanations of various biology concepts had to be repeated in several chapters of the modular biology textbook, adding significantly to its length.

To avoid this drawback, designers may choose to blend the modular approach with other strategies. For example, a hypertext version of *Forty Ways to Save Our Planet* might be authored so that a node explaining greenhouse gases serves as a gateway node to the five dependent nodes.

Modularity via adaptive techniques (AHA!)

AHA! is an adaptive hypermedia educational system well known among researchers. It is adaptive insofar as the content it displays, the quizzing it may conduct, and other system behaviors are personalized to each user's actions. AHA! was created by Paul De Bra, a faculty member at the Technische Universiteit Eindhoven (De Bra, Smits, & Stash, 2006). The most significant application of AHA! has been De Bra's web-based course "Hypermedia Structures and Systems" (http://wwwis.win.tue.nl/2L690), but AHA! functions as a reading system as well as a learning/courseware delivery system.

The unrestricted (or mostly unrestricted) navigation that comes with modularity is central to the design of AHA! AHA!, however, employs digital technology that is far more sophisticated than straightforward hypertext linking. Among the benefits of this technology is the elimination of redundancy. Instead of requiring authors to provide the same foundational information on multiple nodes, AHA! adaptively interposes prerequisite information in response to the user's individual navigation choices. So, as shown below, if the user encounters a mention of Ted Nelson's visionary hypertext system Xanadu without having read the full page explaining it, a capsule explanation of Xanadu is interposed:

> In Xanadu (a fully distributed hypertext system, developed by Ted Nelson at Brown University, from 1965 on) there was only one protocol . . . .
>
> (From the course *Hypermedia Structures and Systems*, De Bra, 2008).

AHA! will then remember that Xanadu has been explained and does not interpose this explanation in the future.

AHA! employs a domain model consisting of the dependency relationships among all the concepts that make up the subject matter. For example, if the domain is algebra, the domain model includes each of the concepts being taught along with a rigorous and formal specification of which concepts should be taught and learned

before which other concepts. At the present time at least, no computer system can build such a domain model. This task must be carried out by the content author.

The second major component of AHA! is a user model. The user model, which is updated continuously as the user navigates through the system, consists first and foremost of a record of the visited nodes. However, AHA! can also question the user and use the results to modify the user model. AHA! can even modify the user model on the basis of assumptions about how much the user may have forgotten between sessions or the user's preferred learning style (e.g., favoring text or visuals) (De Bra, 2008, p. 36). The domain model and the user model jointly make possible the interposition of content elements, such as the explanation of Xanadu, to provide the prerequisite information the user will need as she navigates from one page to another.

In other instances, AHA! makes clear to the user that she lacks the prerequisite information she is likely to need to read certain content elements productively, and offers more suitable alternatives. This feature enables AHA! to manage domains in which complete modularity isn't feasible. The user interface for this feature (which AHA! shares with several other adaptive educational hypermedia systems) is shown in Figure 3. A red dot appears before links to lessons that the user is not prepared for, a green dot before links to lessons that the user can read productively, and a white dot before links to lessons the system thinks the user already knows. When a check mark accompanies a white dot, the user knows the lesson because she has already visited it. If the user is an aggressive reader or learner or has a better background than the system supposes, she can access and read a red dot lesson. Another option in AHA! is for the author to hide or disable links to content the user is not prepared for.

AHA!, then, is a highly sophisticated example of modularity as an approach to designing for selective reading. Later, we will examine SwitchBack, a reading environment that uses much simpler adaptive technology to enhance the summarization approach.

Summarization to support selective reading

Summarization is the final approach to designing documents for selective reading. This approach differs from the others in regard to its purpose. The modularization approach offers readers a choice of which topics to read, and the supported reading pathways approach offers either a choice of which topics to read or a choice in the depth of treatment (e.g., an appendix that expands upon an idea discussed in the body of the document). The summarization approach only offers a choice in depth of treatment, the choice between a summary or the full content.

The most prevalent summarization strategy is the single introductory summary. We often read an introductory summary to decide whether we have any interest in the document. If we do read the full document, we will retain more because we have already read the summary (Lorch & Lorch, 1995). In addition, reading an

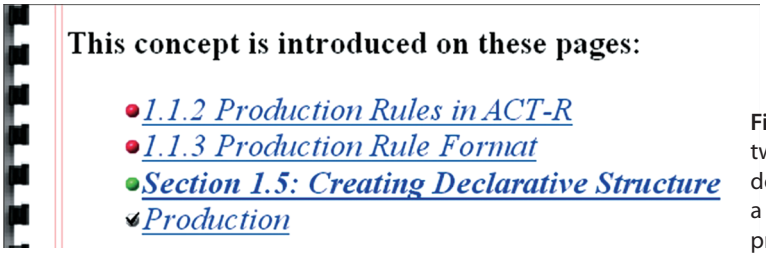


Figure 3. A typical AHA! user interface. The first two links show a red dot, the third shows a green dot, and the fourth shows both a white dot and a check mark indicating that this lesson has previously been visited (De Bra, 2002).

introductory summary offers some support for selective reading. First, the summary provides the reader with a certain amount of prerequisite information, information that substitutes for what the reader misses by reading selectively within the document. Second, because a summary typically maps (at least loosely) the structure of the document, the summary, in conjunction with the document's headings, helps the reader locate sought-for information within the document. Another summarization strategy is the use of multiple summaries. Note that placing a summary at the end of a document or at the end of each chapter promotes retention but does not support selective reading. Now we examine three multiple summary designs.

A BBC website

A web-based historical narrative produced by the BBC (n.d.) concerning the ongoing conflict between India and Pakistan provides us with a straightforward example of the multiple summary strategy. The website consists of an introductory web page followed by nine more pages that make up a chronological sequence. As shown in Figure 4, the reader can choose to read any of the nine chronological pages at two different levels of detail: in

Figure 4. The fourth segment of the BBC narrative. This is the briefest of the nine chronological segments.

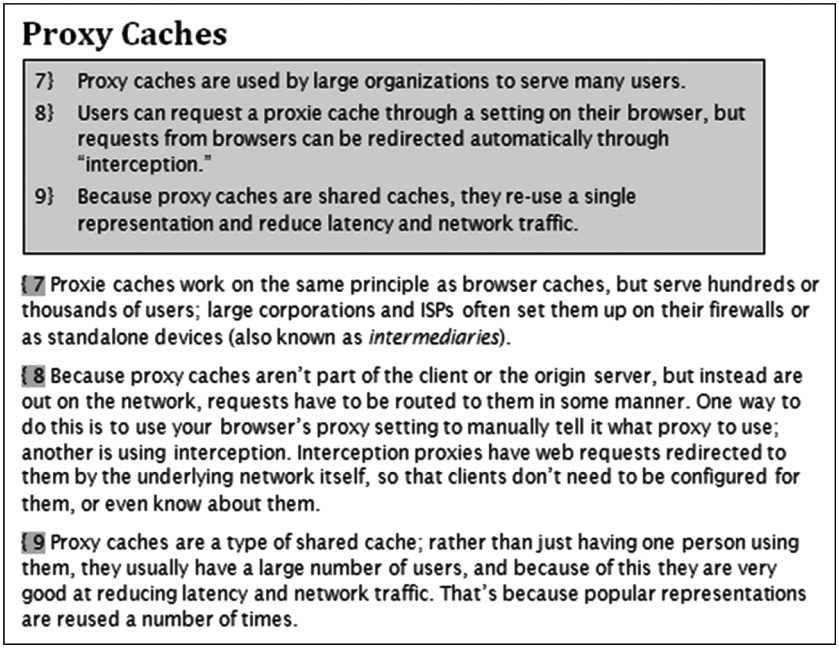
summary or in full. Because each of these summaries explicitly maps to a particular section of the document, the BBC design offers better support for locating information than does a single summary. If you encounter an interesting idea in the summary of the fourth section of the document, you know that the complete discussion appears in the full version of the fourth segment.

This design, however, does not address the problem of lost context. When readers who have read one or more summaries rather than the corresponding sections of the full document switch to the full version of a subsequent section of the document, they may find themselves lacking some relevant background information. The problem of lost context can be reduced by writing longer summaries, but this reduces the savings in time that largely motivates summarization in the first place.

Despite the drawback of lost context, within-document summaries are still effective. Because the reading process is robust, the reader can usually manage even after missing some prerequisite information. If necessary, the reader can hunt backward through the full narratives to find missing information that will alleviate the confusion.

QuikScan

Further possibilities for the multiple summary approach are illustrated by the QuikScan document format (Zhou & Farkas, 2010). QuikScan employs numerous summaries placed strategically within a document, very often directly following a heading. As shown in Figure 5, QuikScan summaries are formatted as numbered list items. These numbers correspond to target numbers placed in the body of the document. Readers can read summaries instead of the detailed content, use the summaries as previews, or navigate quickly to the place where a summarized idea is fully discussed in the body of the document. Empiri-



Proxy Caches

7} Proxy caches are used by large organizations to serve many users.

8} Users can request a proxie cache through a setting on their browser, but requests from browsers can be redirected automatically through "interception."

9} Because proxy caches are shared caches, they re-use a single representation and reduce latency and network traffic.

{ 7 Proxie caches work on the same principle as browser caches, but serve hundreds or thousands of users; large corporations and ISPs often set them up on their firewalls or as standalone devices (also known as *intermediaries*).

{ 8 Because proxy caches aren't part of the client or the origin server, but instead are out on the network, requests have to be routed to them in some manner. One way to do this is to use your browser's proxy setting to manually tell it what proxy to use; another is using interception. Interception proxies have web requests redirected to them by the underlying network itself, so that clients don't need to be configured for them, or even know about them.

{ 9 Proxy caches are a type of shared cache; rather than just having one person using them, they usually have a large number of users, and because of this they are very good at reducing latency and network traffic. That's because popular representations are reused a number of times.

Figure 5. A portion of a QuikScanned document showing a summary following a heading. The three numbered list items in the summary correspond to target numbers in the body of the text.

cal studies demonstrate that reading a QuikScanned document improves comprehension, (Zhou, 2008; Van der Meij & Van der Meij, 2012), facilitates navigation within the document (Zhou & Farkas, 2009), and elicits a positive response from readers (Zhou 2008; Van der Meij & Van der Meij, 2012). It is highly plausible that the increase in comprehension would apply to other instances of the multiple summary strategy.

Because of its numbering scheme, QuikScan allows readers to navigate from a summarized idea to the full discussion at a finer level of granularity than does the BBC design. QuikScan also has potential as an assistive technology for blind readers. In an informal pilot study, a blind student who often waits impatiently while his text-to-speech software pronounces unwanted information jumped directly to the sought-for information (Zhou, 2008). Note, however, that QuikScan, like the BBC website, does not address the lost context problem.

SwitchBack

SwitchBack (2010), shown in Figure 6, is a multiple summary design that does address the lost context problem. It was developed by Farkas and Raleigh at the University of Washington in conjunction with students in the SwitchBack Research Group. It is a working prototype with just basic functionality.

SwitchBack documents are specially authored for the SwitchBack application. The author writes a Lite (summarized) and a Study (complete) version of each section of the document. The reader can then choose whether to read the Lite or Study version of each section. Like AHA!, SwitchBack works adaptively by tracking the

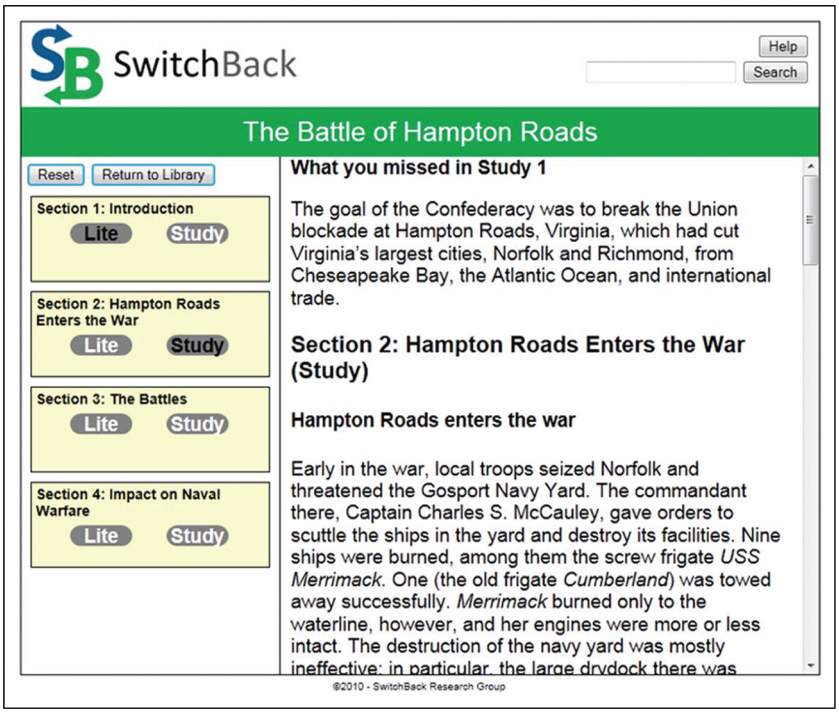


Figure 6. The table of contents of a SwitchBack document and the content displayed when a reader switches from the Lite version of Section 1 to the Study version of Section 2.

reader's navigation through the document and interposing necessary information to prevent a loss of context.

Figure 6 shows a portion of a historical narrative (adapted from Wikipedia) about the Battle of Hampton Roads (often referred to as the battle between the Monitor and Merrimack) during the United States Civil War. It was the first naval battle between ironclad warships. In the figure, we can see that the reader chose to read the Lite version of Section 1 and then switched to the Study version of Section 2. SwitchBack has interposed the prerequisite information (the goal of the Confederacy) that the reader missed by not reading the Study 1 component (the Study version of Section 1). The reader, therefore, is not hindered by the lack of prerequisite information as she reads the Study 2 component. We refer to this interposed prerequisite information as *bridge text* or *bridge components*. If the reader had read Lite 1, Lite 2, and Study 3, SwitchBack would have interposed bridge components 1 and 2 (rather than just bridge component 1). We display bridge components to readers under the heading *What you missed in Study X*. The essence of SwitchBack is simply this: whenever the reader makes a switch from Lite to Study content, any prerequisite information is interposed as bridge text.

Authoring a SwitchBack document requires only a modest technical background, but sophisticated writing skills. Careful decisions must be made about dividing the document into sections and deciding what information belongs in the Lite and bridge components. To make these decisions, the author must keep track of all the pathways readers can take as they choose between the Lite and Study versions of each section. To limit the complexity of this task, the practical limit to the number of sections in a SwitchBack document is five. An MS Word authoring template helps authors keep track of their Lite, Study, and bridge components. While only a working prototype, SwitchBack demonstrates that it is possible to address the lost context problem inherent in multiple summaries by interposing key ideas from sections of the full document that the reader has chosen to bypass.

Conclusion

This survey has identified three broad approaches to designing documents for selective reading and has described a wide range of designs, old and new, simple and complex. These approaches accord with categories of hypertext structures that derive from graph theory (Parunak, 1991; Farkas & Farkas, 2002). The pathways approach corresponds to multipath and hierarchical structures. It constrains navigation and carefully manages dependency relationships. The modularization approach corresponds to the web-like structure (an arbitrary network of nodes). It allows unconstrained navigation, but strives to make each node free-standing because the many incoming links make it difficult to manage dependencies. Finally, multiple summaries are a form of matrix. For example, the navigation panel of *The Battle of Hampton Roads* is a matrix consisting of four rows and two columns. Some matrix hypertexts are written so that each node or cell of the matrix is modular. Multiple summary matrixes, however, are not inherently modular, which is why prerequisite information must often be interposed. One might also choose to regard adaptive techniques as a separate, fourth category. Or, from a somewhat different perspective, regard pathways, modularity, and adaptive techniques as the fundamental categories with summarization, especially multiple summaries, as a hybrid.

In an era in which many people resist reading extended text, longer documents that enable a highly satisfying selective-reading experience are valuable for both individual readers and society as a whole. A future in which such documents are prevalent is better, we think, than an alternative future in which there is less

creation or wide dissemination of longer documents on the grounds that people are not likely to read them. If designers can propose ways to provide a highly satisfying selective-reading experience, perhaps with designs better than those that exist now, it may be possible to strengthen content publishers' commitment to extended text.

References

Albers, M. J. (2003). Multidimensional audience analysis for dynamic information. *Journal of Technical Writing and Communication, 33*(3), 263–79.

Anderson, R. C. (2004). Role of the reader's schema in comprehension, learning, and memory. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed.) (pp. 594–606). Newark, DE: International Reading Association.

BBC. (n.d.). India-Pakistan: Troubled relations. Retrieved 2/17/13, from http://news.bbc.co.uk/hi/english/static/in_depth/south_asia/2002/india_pakistan/timeline/default.stm

Boiko, B. (2005). *Content management bible* (2nd ed.). Indianapolis, IN: Wiley Publishing.

Byliner (2012). Retrieved 2/17/13, from https://www.byliner.com.

Carr, N. (2010). *The shallows: What the Internet is doing to our brains*. New York: Norton.

Chambliss, M. J., & Calfee, R. C. (1989). Designing science textbooks to enhance student understanding. *Educational Psychologist, 24*(3), 307–322.

De Bra, P. (2008). Adaptive hypermedia. In H. H. Adelsberger, P. Kinshuk, J. M. Pawlowski, & D. Sampson (Eds.), *Handbook on information technologies for education and training* (pp. 29-46). Berlin: Springer-Verlag.

De Bra, P. (2002). Adaptive educational hypermedia on the Web. *Communications of the ACM, 45*(5), 60–61.

De Bra, P., Smits, D., & Stash, N. (2006). Creating and delivering adaptive courses with AHA! Proceedings of *The First European Conference on Technology Enhanced Learning*, EC-TEL, Springer LNCS 4227, pp. 21–33, Crete, October 1–4.

Dignan, L. (2010). *Amazon's Kindle Singles e-books: One shrewd business move.* Retrieved 2/17/13,from http://www.zdnet.com/blog/btl/amazons-kindle-singles-e-books-one-shrewd-business-move/40320

Eppler, M. J., & Mengis, J. M. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society: An International Journal, 20*(5), 325–344.

Farkas, D. K., & Farkas, J. B. (2002). *Principles of web design*. New York, NY: Longman.

Graham, G. (2009). *How to write a white paper for multiple audiences.* Retrieved 2/17/13, from http://www.thatwhitepaperguy.com/white-paper-article-how-to-write-a-white-paper-for-multiple-audiences.html

Holland, V. M., Charrow, V. R., & Wright, W. W. (1988). How can technical writers write effectively for several audiences at once? In L. Beene, & P. White (Eds.), *Solving problems in technical writing* (pp. 27–54). Oxford: Oxford University Press.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323.

Lee, M. J. (2005). Expanding hypertext: Does it address disorientation? Depends on individuals' adventurousness. *Journal of Computer-Mediated Communication, 10*(3). Retrieved 2/17/13, from http://jcmc.indiana.edu/vol10/issue3/lee.html

Liu, Z. (2005). Reading behavior in the digital environment. *Journal of Documentation, 61*(6), 700–712.

Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review, 1*(3), 209–234.

Lorch, R. F., & Lorch, E. P. (1995). Effects of organizational signals on text processing strategies. *Journal of Educational Psychology, 87*(4), 537–544.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes, 22*(3), 247–288.

Menzies, H., & Newson, J. (2007). No time to think: Academics' life in the globally wired university. *Time and Society, 16*(1), 83–98.

Nicholas, D., Huntington, P., Jamali, H. R., Rowlands, I., Dobrowolski, T., & Tenopir, C. (2008). Viewing and reading behaviour in a virtual environment: The full-text download and what can be read into it. Proceedings of the *Association for Information Management, 60*, 185–198.

Palfrey, J., & Gasser, U. (2008). *Born digital.* New York: Basic Books.

Parunak, H. V. D. (1991). Ordering the information graph. In J.

Devlin, & E. Berk (Eds.), *Hypertext/ hypermedia handbook* (pp. 299–325). New York: McGraw-Hill.

Richtel, M. (2010). *Growing up digital, wired for distraction.* Retrieved 2/17/13, from http://www.nytimes.com/2010/11/21/technology/21brain.html?_r=4&pagewante+d=all

Rockley, A. (2001). Dynamic content management. *Intercom*, February 2001, 28–32, and 42.

Schick, A. G., Gordon, L. A., & Haka, S. (1990). Information overload: A temporal approach. *Accounting, Organizations and Society, 15*(3), 199–220.

Severson, E. (2009). *DITA and dynamic content delivery*. Retrieved 2/17/13, from http://www.flatironssolutions.com/blog/dita-and-dynamic-content-delivery/

Springer. (n.d.). *Introducing SpringerBriefs*. Retrieved 2/17/13, from http://www.springer.com/authors/book+authors/springerbriefs?SGWID=0-1720013-0-0-0

SwitchBack. (2010). *SwitchBack Research Group web page.* Retrieved2/17/13, from http://faculty.washington.edu/farkas

Thomas, G. F., King, C. L., Cook, L., Keiterman, M., Miller, S., & Wardle, A. (2006). Reconceptualizing e-mail overload. *Journal of Business and Technical Communication, 20*(3), 252–287.

Van der Meij. H., & Van der Meij, J. (2012). Improving text recall with multiple summaries. *British Journal of Educational Psychology, 82*(2), 257–269.

Wurman. R. S. (2001). *Information anxiety 2*. Indianapolis, IN: Que.

Zhou, Q. (2008). *QuikScan: facilitating document use through innovative formatting*. Seattle, Washington. Unpublished doctoral dissertation, Department of Human Centered Design & Engineering, University of Washington.

Zhou, Q., & Farkas, D. K. (2010). QuikScan: Formatting documents for better comprehension and navigation. *Technical Communication, 57*(2), 197–209. See also www.QuikScan.org.

Zhou, Q., & Farkas, D. K. (2009). QuikScan: Facilitating reading and information navigation through innovative document formatting. Proceedings of the *International Professional Communication Conference,* Honolulu, Hawaii.

About the authors

David K. Farkas (http://faculty.washington.edu/farkas) is a Professor specializing in information design in the Department of Human Centered Design & Engineering at the University of Washington. His current research areas are software user assistance, consumer-information graphics, slideware design, and reading environments (QuikScan, SwitchBack) that accommodate changing reading habits by supporting selective reading. He is a Fellow of the Society for Technical Communication and a recipient of STC's Jay R. Gould Award for excellence in teaching.

Email: farkas@uw.edu

Christopher A. Raleigh is a technical writer for the Intel Corporation, specializing in software documentation. He holds a Bachelor of Arts in English from Eastern Washington University and a Master of Science in Technical Communication from the University of Washington, where he focused on user-centered design and educational research. Christopher has worked as a proposal coordinator, instructor for the University of Washington College of Engineering, research analyst, and technical writer/editor. He has a broad range of software and web development, instructional design, and user-interface design skills.

Email: craleigh@uw.edu

Contact

David K. Farkas
University of Washington
Department of Human Centered Design & Engineering
College of Engineering, Sieg Hall
Box 352315
Seattle, Washington, USA, 98195

Christopher A. Raleigh
PO Box 6556
Portland, Oregon, USA 97228-6556