University of Washington, Seattle, Washington
Spring 1998

Web Design

An Empiricist's Guide

By Mary Evans

A Project Report
submitted in partial fulfillment of the requirements for
Master of Science in Technical Communication

Table of Contents

Acknowledgments

I am especially grateful to the members of my committee, Professor Judy Ramey, Committee Chair, and Professors Mary Coney and Thomas Williams, for cheerful and expert help with research, thinking, and writing, and insistence on careful scholarship. I also would like to express my appreciation to the other faculty members of the Department of Technical Communication with whom I studied, and to the Department's staff members, who helped me to avoid any missteps on my path towards my degree. As I leave the Department, I take with me many happy memories of my time as a Technical Communication student, confidence in the excellence of the training I received there, and gratitude for the encouragement and support that I received from all members of the faculty and staff.

Introduction

The World Wide Web can be a powerful tool for government and non-profit agencies, whose mission is to offer services and information at low cost to a wide variety of people. Not only can the Web help us to achieve this end in an unprecedentedly effective way, but it also eventually may prove to be a way to share resources with people who are often under-served, especially people living far from service centers and citizens of developing countries.

Before our web sites can achieve their full potential, however, we need effective site design techniques. But while empirical research and tradition offer reliable guidelines for preparing print works, even the most fundamental web design questions remain open. For example,

- How should web sites be structured? Hierarchically-structured sites are common, but is a hierarchy the best structure?
- How can links be designed to make their destinations clear to readers?
- Should links be arranged vertically along the side of a web page, horizontally across the top or bottom, embedded in body text, or presented in some other configuration?

Shneiderman (1997) commented that "It will take a decade until sufficient experience, experimentation, and hypothesis testing clarify issues," and warned that meanwhile, "the paucity of empirical data to validate or sharpen insight means that some guidelines are misleading."

Nevertheless, many sets of web design guidelines have been published. A list of some of the best-known ones is included in the appendix to this report. But the basis of these guidelines, and the degree to which they may have been derived from empirical research are not clear.

My aim in preparing this report was to review empirical research results in an attempt to find research-based answers to key questions of web site design. In this report, I review the results of relevant empirical research, summarize those results, and then indicate what those results seem to suggest to people trying to design effective web sites. I also have tried to indicate areas where further research into particular issues would be helpful to the web design community.

Important caveats are that

- In the cases of most web design questions, few empirical findings are available to us. Too often, we can't know whether the results from one study would be replicated across experiments, simply because no like studies have been performed.
- Most of the findings reviewed in this report were collected by studying systems that are similar but not identical to web sites. Some of these studies were conducted in controlled settings very different from those in which most web users work. Whether these findings can be generalized to the web is often impossible to determine.

For these reasons, at least some generalizations I have drawn from my review of these studies may eventually be proved wrong by new information.

Why, then, have I chosen an empirical approach? An obvious response to this question is to note that, even with only scarce information on which to base design decisions, and even considering the possibility that designers could be misled by some of that information, it's nevertheless necessary to base decisions on something. But a better reason is that empirical research can "ground-truth" other kinds of design knowledge, such as intuition derived from personal experience. Designers using any other approach alone, in the absence of empirical knowledge, can too easily be fooled. Norman (1991) reviewed studies of people using computer menu systems, and commented that "Although intuition aids in good design, there are many cases in which the best design is initially counterintuitive even to experts in the field."[1]

---

[1] Norman went on to comment that, although it's all too easy to write guidelines based only on intuition or experience, "It may be even more dangerous to write them armed with only a little knowledge." Perhaps, then, the best-equipped web designers are those who bring to their work a combination of both intuition and empiricism.

Web Usability

My emphasis in this report is on designing web sites to be as **usable** as possible.

What makes a web site usable? Nielsen (1993), who has extensively studied people using hypertext systems and the problems they encounter, has proposed five usability attributes for hypertext systems. Applying his approach to the web—which is, after all, the best-known hypertext system—we see that a web site would be judged to be usable if it has these attributes:

- **easy to learn** (people can readily see how the site is structured and how to navigate through it),
- **efficient to use** (once someone arrives at a page in the site, she can easily see its relation to the rest of the site),
- **easy to remember** (after a period of nonuse, people can still remember how to navigate in the site),[2]
- **few errors** (people will "rarely follow a link only to discover that they really did not want to go to wherever the link leads"),
- **pleasant to use** (people are pleased with what they find at a web site).

Of course, usability is not the only criterion for an effective web site. Nielsen also noted that the fundamental issue is really whether a product is **useful**: that is, "whether the system can be used to achieve some desired goal." To Nielsen, usefulness incorporates both **utility** (whether the system has the functionality to achieve a goal) and **usability** (whether people can use the system to achieve a goal). For a given web site, it is the quality and completeness of the site's content, and whether that content can support the tasks performed by visitors to the site that determine its utility. Its usability is determined by whether visitors can find and use that content easily and effectively—and by whether their visits to the site prove pleasant or frustrating. In this report, I have focused on how

[2] The Web differs from other devices and software in that people are likely to visit most sites only once or intermittently. They are unlikely to visit most sites regularly in the way they might regularly use a word-processing application or copier. Hence, for the Web, "easy to remember" is likely less important than "easy to learn."

best to achieve usable design, leaving the question of how to ensure the utility of web sites up to their designers.

What kinds of research studies are relevant for web design?

Although very few studies of web site usability have been completed, the results of research into the usability of other kinds of human-computer interfaces and interface features offer web designers clues about which design approaches may be better or worse. Relevant research includes

- studies of human cognition and perception.
- studies of hypermedia, screen design, direct manipulation, and menu selection.
- research on web usability (however, very few of these studies have yet been completed).

To answer the web design questions discussed in this report, I have examined studies in all these categories.

Web Design Questions

On the following pages are discussions of key web design questions that fall within five broad subject categories:

1. **Designing for Users' Technology**—managing the effects of system response time.
2. **Preparing Content for the Web**—designing readable documents and choosing words meaningful to readers.
3. **Organizing a Web Site**—choosing an overall topology and number of levels.
4. **Designing Navigation Controls**—wording and arranging links, implementing text and image links, and designing information overviews.
5. **Designing Web Pages**—choosing the amount of white space, ensuring text legibility, deciding on text and background colors, and using animations and the BLINK tag.

Designing for Users' Technology

Page Downloading Time

Q. How can designers minimize page downloading times? And how important is this issue?

A. Surveys of web users have indicated that most people who connect to the web do so using modems, and hence have relatively slow connection speeds. Respondents have consistently identified slow page downloading as an important problem with the web. Perhaps more important than users' feelings of dissatisfaction with slow page downloading, however, is the effect of delays in system response on users' task performance. In one study, people using a computer system made more errors when there were delays of more than 12 seconds in the system's response to their actions. This results has been interpreted to mean that long delays may strain users' working memory, impairing their ability to complete tasks. It is possible that delays in page downloading may produce a similar effect.

Page downloading speed is determined primarily by the size of the page and the connection speed at which it is being downloaded. Web site designers usually can control only the sizes of pages at their sites; page files that are smaller in size can be downloaded faster.

What can web designers do?

- Because government should offer universal access to services, as much as possible, government site developers should, as a general rule, design for people with only modem access to the web. (Agencies with extensive web development resources may wish to design more than one version of a site: a version that can be universally accessed, and a version that incorporates recent technology.)
- To reduce page downloading time, avoid large graphics and, generally, use few graphics.
- Avoid graphics that are purely ornamental; instead, use graphics that reinforce associated text information on the page.

- Reduce the number of colors in graphic files when possible, and use the highest compression ratio possible when creating jpeg images.
- Provide text link equivalents for all graphics that operate as link hot spots or image maps, to serve people working with graphics off, as well as people using text-only browsers (such as Lynx) or screen readers.
- Include height and width attributes in image tags, as follows: "<IMG height=yy width=xx SRC=image.gif" (here, pixels are the units of height and width). Incorporating these tags speeds the display of graphics, because the viewer's browser does not have to compute the dimensions of the bounding box that will contain the image (Sano 1996).
- As recommended by Waters (1997), use the 216-color "browser-safe" color palette in your graphics application when you create graphics for the web (this palette can be downloaded from www.lynda.com/hex.html). These colors look identical on Macintosh computers and in Microsoft Windows. When you use colors not in this palette, a browser loading the image will mix pixels of closely-related colors in an attempt to match the colors you chose. This color mixing is called dithering, and it requires processing time. As Waters has noted, a dithered image is generally larger than one that is not dithered.

See Also

Choosing the Number of Levels in the Site 26

Backgrounder

System Response Times on the Web The Graphics, Visualization, and Usability Center (GVU Center) at the Georgia Institute of Technology has surveyed web users approximately every 6 months since 1994. The most recent surveys have been advertised by Yahoo! (yahoo.com), Netscape (home.netscape.com), and WebTV. Because people voluntarily fill out the survey, respondents are self-selected rather than randomly chosen. Hence, it is not possible to know the degree to which the survey results are representative of the web user community as a whole. However, the GVU surveys remain the most extensive public-domain survey of web users.

The GVU Center's surveys of web users have consistently shown that

- There is a great disparity among respondents in the speed with which they are able to access the web: from home users with 14.4 or 28.8 kbps modems to users at large corporations, government agency headquarters, and universities with T1 or faster access.

- Most respondents connect to the web by modem, the slowest kind of web connection. Results from the most recent survey are shown in Table 1, below: of the 7,695 respondents, about 60 percent were using modems; 66 percent had been using modems during the immediately previous survey, run during April and May, 1997. The survey team reported that "...the difference in percentages is mainly accounted for by those who are unsure of their connection speed" (GVU Center 1997).

Table 1. Percentages of 7,695 respondents to the GVU Center's Eighth WWW Survey (conducted during October and November 1997) who reported connecting to the web at various connection speeds.

Connection Speed	Percentage of Respondents
>45 Mb/sec (FDDI)	0.8
45 Mb/sec (T3)	1.6
10 Mb/sec	2.3
1 Mb/sec (T1)	9.1
128 Kb/sec (ISDN)	4.2
56 Kb/sec	6.1
33 Kb/sec (modem)	27.2
28 Kb/sec (modem)	21.7
14 Kb/sec (modem)	5.4
<14 Kb/sec	0.4
Don't know	21.3

The average modem speed increased somewhat between the immediately previous and current survey: 27 percent of the 7,695 respondents to the most recent survey reported that they use 33.6K modems, compared with 20 percent of the respondents to the previous (April-May 1997) survey. If these respondents are representative of the web user population as a whole, these results suggest that although web connection speeds are gradually increasing, most web users still have relatively slow connections. Unless the rate of improvement in connection speed increases substantially, most web users are likely to continue to have connections substantially slower than the maximum speeds now possible during the next few years. These survey results suggest that web sites intended for audiences typical of the web user community should be designed for people using relatively slow connections.

The GVU Center's survey has also consistently shown that most respondents (63 percent in the most recent survey, as shown in Table 2) consider slow page downloading times to be an important problem with the web.

Table 2. Percentages of 7,220 survey respondents who indicated that the following issues are among the most important problems with the web. (Respondents were offered only the choices shown; they could not indicate other choices.)

Problem	Percentage of Respondents
It takes too long to view/download pages	62.8
Encountering links that do not work ("linkrot")	59.7
Not being able to find information I'm looking for	49.5
Not being able to find a page I know is out there	32.3
Not being able to efficiently organize the information I gather	25.3
Not being able to return to a page I once visited	17.8
Other	9.5
Not being able to visualize where I am and where I can go	8.8
Not being able to determine where I am	6.4
It costs too much	6.0

Many survey respondents have upgraded their connection speed, or plan to soon (Table 3). Again, this finding indicates that users feel a strong need to improve their connection speed, and that speed is an important issue for web users.

Table 3. Percentages of respondents answering the question, "Has the speed of your primary connection to the Internet been upgraded (made faster) in the past year?"

Speed Upgrade in Past Year?	Percentage of Respondents
Yes, and plan to upgrade again within 6 months	16.9
Yes, and do not plan to upgrade again within 6 months	21.8
No, but plan to upgrade in the next 6 months	14.1
No, and do not plan to upgrade in the next 6 months	23.5
Don't know	23.8

Despite the generally slow connection speeds that prevail among respondents, the great majority—90 percent in the current survey—choose to automatically download images most of the time. The survey team commented that this, along with generally slow connection speeds of most users "probably explains why downloading speed is an issue for most users."

Effects of System Response Time Page downloading time is an example of what researchers have termed system response time (SRT), which is the time taken by a computer system to respond to user actions such as keystrokes and cursor movements. How might the relatively slow system response times typical of the web affect web users?

Patterson and Egido (1986) reviewed previous research on the effects of system response time on computer system users, and concluded that "In general, there is agreement in the literature about users not liking long SRTs." However, the effect of system response time on users' performance does not appear to be straightforward.

Mayhew (1991) suggested that people may do their work differently depending on whether system response time is fast or slow, making more frequent use of trial and error when response times are fast, and using more deliberate strategies when the system responds more slowly. Patterson and Egido (1986) obtained evidence supporting this idea. They observed 16 women who were given print pictures and then searched for the same pictures in a database, when the system response time was fast (3 seconds per picture) and slow (11 seconds per picture). When the system responded slowly, the test subjects viewed fewer pictures before making each decision (that is, before selecting a particular picture as the correct match). This result suggests that to avoid delays, the subjects may have observed each picture more carefully when the system responded more slowly.

However, although people may adapt their work strategies to adjust for system response time, there is evidence that especially long response times may degrade users' ability to complete tasks, perhaps by straining working memory. During a 4-day experiment, Barber and Lucas (1983) observed 19 circuit layout clerks using terminals on which system response time had been slowed to an average of about 14 seconds, and compared their performance with that of 81 other clerks using terminals with average response times of about 6 seconds (all these circuit layout clerks managed the connection and disconnection of telephone circuits). The authors found that error rates were highest when the system responded either fastest or most slowly to user actions; the fewest errors were made when system response time was about 12 seconds. They suggested that as system response time increased, clerks became more careful when entering transactions in order to avoid repeating transactions—hence reducing their error rate from the higher rates observed at the fastest system response times—but as response time increased beyond 12 seconds, working memory was strained, making errors more likely as the clerks lost track of their place in their tasks.

Miller (1968), in an early theoretical paper on the effects of response time, proposed that the length of system response delays that users can tolerate depends on the tasks they are performing. Because no studies have yet been performed of the effects of page downloading times on web users' performance, web site designers cannot know just what times may cause problems for web users. In the absence of this information, the best design strategy seems to be simply to reduce page downloading times as much as possible.

Preparing Content for the Web

Choosing the Site's Vocabulary

Q. How important is it to tailor the vocabulary used in a web site to its users? If it is important, how can designers do it?

A. For many years, technical communicators have stressed the need to use language that's meaningful to readers. That this would be helpful to people seems intuitively obvious. However, a difficulty in accomplishing this may be less obvious: people differ widely in the terms they choose to describe particular concepts**.** Because of this variability,

- There may be times when no terms are meaningful to all users of a web site.
- It's often impossible to choose a single term for an information object or system function that's meaningful to all people.

What can web designers do?

- Make ensuring the usability of language a priority goal of site usability tests.
- Use field studies, user interviews, card sorting, and/or other audience analysis techniques or participatory design to discover and implement users' vocabularies; see Wixon and Ramey (1996) for descriptions of some of these methods.
- In sites designed for multiple audiences, use the preceding methods to tailor the terminology in sections of the site designed for different audiences to fit each of those audiences.
- "Richly index" site search engines. That is, include multiple keywords for each information object so that one or another keyword will be meaningful to each person searching the site. Furnas (1985) suggested an intriguing method for creating these multiple keywords, called "adaptive indexing," that may be especially suitable for web design (see below).

- Whenever space allows, include more information about each link, in the form of link abstracts (a short abstract with each link explaining what users will see if they click that link), to help users browsing through a site to find information. (See “Making Links Predictive” for more details on how to do this.)

See Also

Writing to be Read Online 16

Making Links Predictable 33

Backgrounder

The issue of using users’ language—which we might call language usability—is important in web design for three reasons:

- It’s important to use users’ language in order to make the site’s content understandable.
- Language usability is a menu and link design issue: what do you call options and controls on a screen so that people recognize them?
- Site search engines match stored keywords to user-entered search criteria in order to retrieve information objects from the site. When these keywords differ from users’ search criteria, no match will be retrieved, even when the information object the user wants exists in the site.

Technical communicators agree that it’s important to use language familiar to users in system interfaces and documentation. Even early observers of computer interfaces recognized the need to match system and users’ vocabularies. Chapanis (1965) maintained that “changes in the words used in man-machine systems may produce greater improvements in performance than human engineering changes in the machine itself.” He distinguished between the **readability** of text—whether a text can be easily read because the words it contains are known to its reader—and its **intelligibility**—whether the message the writer hoped to encode in the text is transmitted intact to the reader.

But only recently have researchers made quantitative studies of the effects of terminology choices on the ability of software users to complete tasks. Most of these studies have focused on the effects of different choices of search keywords and command names on users' success at retrieving desired information or initiating system actions.

One study focused on the ability of users to follow instructions written in very technical language. Bloom (1987-88) observed users as they attempted to follow the complex instructions for an unfamiliar word-processing application, a task that he maintained can cause "information overload." This overload, Bloom contended, can cause even the most motivated and intelligent users to fail to complete tasks. He referred to earlier studies showing that users often become confused because they invoke misleading associations for unfamiliar technical terms. Bloom found that when task instructions contained unfamiliar technical jargon, users were eventually able to follow the instructions to complete a task, although they made more errors than users following instructions written in familiar language. However, those users following jargon-laden instructions were less able to generalize from their experience in order to complete new, related tasks. From these results, Bloom concluded that unfamiliar technical terminology reduces users' comprehension of instructions, although it does not necessarily prevent them from eventually learning to follow instructions.

Bloom's results appear relevant wherever instructions are presented to users. In web sites, that would be on pages presenting forms to fill out and submit, and on pages displaying instructions for downloading items, performing searches, or running applets. Bloom did not specifically test the effects of technical language on users' comprehension of expository texts. However, it also seems reasonable to infer from his study that avoiding unnecessarily technical language in any kind of text would improve comprehension.

Furnas et al. (1984 and 1987), in a series of experiments both in the laboratory and with people using actual systems, examined what they term the "vocabulary problem": new or intermittent users often cannot obtain the actions or information that they want from a system because they do not know the right words for making their requests. These authors observed that people use a great variety of words to refer to the same thing or concept. They found, for example, that the probability that two typists will choose the same name for an editing function in a software program is less than one in 14. They also found that requiring users to type in a single correct word in order to access a system function—a common interface design practice—caused failure rates of 80 to 90 percent. They concluded that "The idea of an 'obvious,' 'self-evident,' or 'natural' term is a myth," and recommended offering more than one access term to users trying to find information objects or system actions.

Remde et al. (1987) reviewed earlier studies of people's problems using information retrieval interfaces, and concluded that the main reason users often fail to find information that is present in the system is that "users describe the things they want in different terms from those by which the system knows them." They found that "greatly increasing the number of names by which each information object can be reached"—a technique known as "rich indexing" or "unlimited aliasing"—can increase search success rates by 20 to 80 percent, presumably because in a sufficiently richly indexed interface, at least one term intuitive for each user is available.

Furnas (1985) suggested using a technique he termed "adaptive indexing" to discover the keywords that a system's real users would choose as their search criteria. An adaptive indexing system would use real keyword retrieval activity to identify the words users choose as search criteria. Furnas gave a hypothetical example of a user of a manual searching for information about the print function, who might first type in "hardcopy," then "output," then "type," and then finally "print." "Print" is a keyword that the system recognizes, so it provides the information to the user. It also then identifies the previously-attempted keywords ("hardcopy," "output," and "type") as this user's synonyms for "print," and includes those terms in its list of keywords for accessing information about the print function. Capturing users' keywords in this way might aid designers of search engines for large web sites.

Perhaps in only one study have researchers directly examined the usability of the vocabulary used in web sites. Spool et al. (1997) observed more than 50 people searching for information in nine commercial web sites. They commented that some of these web sites assumed more domain knowledge than users had, and that these sites tended to get low usability ratings from test subjects. However, they did not offer any quantitative measure of this effect. They described one instance in which test subjects had difficulty with terminology: at the Travelocity site (www.travelocity.com), using "segments" instead of "round trips" confused participants trying to compare the costs of alternative flights between two cities. These findings are in accord with those of earlier researchers, adding weight to the idea that taking care to adapt a web site's vocabulary to its users will make it more usable.

Writing to be Read Online

Q. How should text written for online reading differ from print texts? (For example, should it be equally or more concise?)

A. Most people read online text differently from how they read printed texts. Rather than reading word by word, most people quickly scan blocks of online text. When reading text on pages within a web site, most people also move quickly among pages.

What can web designers do?

Morkes and Nielsen (1997 and 1998), after completing a series of studies of people reading text on web sites, recommended that text designed to fit this markedly different reading style should have the following attributes:

- **It should be easy to scan.** Ways to improve scannability suggested by Morkes and Nielsen include adding tables of content and section summaries, using bulleted and numbered lists, highlighting key words and phrases, adding meaningful headings and subheadings, and discussing only one main idea in each paragraph.
- **It should be concise**, with language streamlined and non-essential details removed.
- **It should be objective in tone.** A surprising result of Morkes and Nielsen's study was that self-promotional language not only was disliked by test subjects, but also adversely affected usability. Using an objective tone improved usability (and was preferred by subjects). For this reason, Morkes and Nielsen suggest removing adjectives such as "great" or "overwhelming," jargon words such as "paradigm," and claims that are not supported by evidence.

Not all texts that might be included in an informational website can be so extensively rewritten. For example, many agency websites offer online versions of legislation and other formal public documents. A good guideline might be to revise all texts intended for online reading according to Morkes and Nielsen's recommendations, and to offer other documents in forms that can be readily printed out (such as PDF, Portable Document Format).

See Also

Backgrounder

Several observers have noted that most people read online texts in a very different way than they read print texts. In an extensive study of how 71 users read text in a web site, Morkes and Nielsen (1997) found that "79 percent of our test users always scanned any new page they came across; only 16 percent read word-by-word." Bachiochi et al. (1997), in a study of how 40 Aetna employees navigated within a 45-page web site, observed that "Users do not read blocks of text; they jump around and react to links." Spool et al. (1997) observed more than 50 people searching for information in nine commercial web sites, and reported that generally, test participants skimmed pages, looking for relevant information; only when they found an item of interest did they read text completely. These authors reported that page designs that facilitated skimming "fared best" in their tests (but did not quantify their findings).

Not only do most readers appear to skim rather than read online texts, but they also may be less likely than readers of print documents to read or skim completely through all the available information about a topic. Hardman (1989) observed people using Glasgow Online, a HyperCard city guide to Glasgow, Scotland, to complete tasks representative of information searches visitors to that city might make. She found that people typically were reluctant to search thoroughly for information. For example, once they had identified one or two hotels that met their criteria, they stopped viewing hotel information, even though other hotels might also have been acceptable.

Recently, a research team at Sun Microsystems attempted to identify the attributes of online text that make it more or less readable. Morkes and Nielsen (1997) tested four different versions of a small web site, "Travel Nebraska," with 41 users who were not computer professionals or writers. The four versions of the site included

(1) a control version written in promotional language (which the authors view as characteristic of many web sites),

(2) a concise version, also written in promotional language, but containing much shorter text than the original,

(3) a “scannable” version, also written in promotional language, which incorporated bulleted lists, bolding of key sentences and phrases, photo captions, shorter text sections, and more headings, and

(4) an objective version, which “presented information without exaggeration, subjective claims, or boasting.”[3]

Each test subject was asked to

(1) use one of the four versions to complete two information-searching tasks and a judgment task,

(2) fill out a subjective evaluation questionnaire,

(3) take an exam testing knowledge of the site’s contents, and

(4) draw a diagram of the site structure.

The authors constructed a “usability score” for each site version, which incorporated the average time taken to complete tasks, the number of errors made, subjects’ memory for site contents (measured by both recall and recognition tests), subjects’ subjective satisfaction, the time taken to complete a site diagram, and the accuracy of site diagrams. They found that “Overall usability scores for all versions of the site show that, compared to the control version [the promotional version], the scannable version is 47% better, the concise version 58% better, and the objective version 27% better.” The authors concluded that writing texts that are concise, scannable, and objective in tone will “make a positive difference in Web users’ performance and subjective satisfaction.” Nielsen (1997b), commenting on this study, speculated that “promotional language imposes a cognitive burden on users who have to filter out the hyperbole to get at the facts.”

Later, Morkes and Nielsen (1998) tested their conclusion that concise, scannable, and objective text is best for the web. They created two versions of another study website. The contents of each version were derived from two whitepapers from the Sun Microsystems site. The “original” version contained three pages (a home page and a page containing each whitepaper), and only slightly modified texts from the whitepapers. The

[3] For example, while the first sentence in the control version was, “Nebraska is filled with internationally recognized attractions that draw large crowds of people every year, without fail,” the first sentence in the objective version was, “Nebraska has several attractions.”

"rewritten" version contained eight pages, all much shorter than the pages of the original version, and texts rewritten to be

- concise (the texts contained 54 percent as many words as the original whitepapers),
- scannable (the authors shortened paragraphs and added tables of contents, section summaries, bulleted and numbered lists, highlighted keywords, and more headings), and
- objective (the authors removed value-laden adjectives such as "great" or "overwhelming," buzzwords such as "paradigm," and claims that were not supported by evidence).

They asked 21 technical users to use one or the other version to perform information-searching and judgment tasks, to take an exam measuring knowledge of the site contents, and to complete a questionnaire measuring subjective satisfaction with the site. The authors then computed an overall usability score for each site version from measurements of time taken to complete tasks, number of errors made, memory for the contents of the site, and subjective satisfaction. They found that "The rewritten website scored 159% higher than the original in measured usability." (That is, the usability score of the original site was 100, and that of the rewritten site was 159.)

Organizing a Web Site

Site Topology

Q. What's the best way to organize information in a web site? Hierarchical sites are the most common, but is a hierarchy always the best structure?

A. A difficulty for web site designers is that people appear to differ widely in how they navigate through hypertext documents or web sites. Generally, web users seem to be either (1) searchers, who search a site in a very directed way for a particular piece of information of interest to them, or (2) browsers, who take a more open-ended, exploratory approach to navigation.

Researchers have proposed that a hierarchical site structure better supports information searching, while a network structure better supports browsing. Indeed, in one study, people seemed to be better able to find information in hierarchical sites than in sites organized as networks. But these test subjects were able to find information fastest in sites organized as a combination of hierarchical and network structure.

Perhaps the fundamental principle is that it's important to arrange information within a web site in a way that is meaningful to users: in another study, people were able to find information in a database fastest after users had participated in organizing the information in that database.

What can web designers do?

- For informational web sites, choose either a hierarchical structure or a structure that is primarily hierarchical but that also includes network-style links among pages containing related information (even though those pages may not be structurally close to each other).
- Use card-sorting (as described below) to arrange items in a web site naturally. However, note that this approach is likely to work best if there's just one audience for the site.

See Also

Information Overviews 57

Choosing the Number of Levels in the Site 26

Backgrounder

Hypertext documents are most often arranged hierarchically, with more general concepts and information at the highest levels of the hierarchy and more specific concepts or information at lower levels. However, it's possible to organize web sites in other ways. A site organized according to a linear structure would be navigated sequentially, like a printed document. At the opposite end of the site topology continuum from hierarchical sites would be sites organized as pure networks. Many web sites are organized as a combination of hierarchy and network. Such sites are organized into distinct levels, with more general information at higher levels, but with pages at different levels linked when they contain related information.

For web designers, there is a key question: Is a hierarchical structure best for web sites, or would another structure be better? To answer this question, it's necessary to examine how users navigate within web sites or hypertext documents.

Users' Navigation Styles Many researchers have noted that hypertext users differ in their general approach to navigation, some preferring to search for specific information, and others displaying a more open-ended navigation style. For example, Ehrlich, in Perlman (1990), commented that "Based on previous research we identified two major styles of navigation. One style, called browsing, is characterized by the user who searches through a large number of options until an appropriate topic is recognized. The other style, called analytical search, is characterized by the user who generates a short list of options based on recall of the topic."

Users with different navigation styles use features of a hypertext document or web site differently. For example, Hammond and Allinson (as described in Nielsen 1989) found that users exploring a hypertext system in order to learn about a subject made little use of the system's index but heavily used the system's guided tour feature. In contrast, users trying to answer specific questions used the index much more heavily than the guided tour feature.

Users with different navigation styles also may differ in their preferences for different site structures. Smith et al. (1997), after reviewing the results of research in hypertext navigation, concluded that "...exploratory tasks are best supported by a network or combination information structure, whilst searching tasks are best supported by a hierarchical information structure."

A Comparison of Topologies Mohageg (1992) compared four different hypertext linking structures—linear, network (only relational links, represented by keyword links), hierarchical (only structural links), and combination (hierarchical and network)—in two experiments in which 62 hypertext novices performed information retrieval tasks using a geography database in HyperCard. He measured task completion times, number of errors, deviation from the shortest possible path to take to complete each task, and uncertainty (a measure that reflects the degree of difference among the paths chosen by individual subjects completing the same task).

In Mohageg's Experiment A, each hypertext structure was arranged so that subjects would need to traverse the same number of links to complete each task, regardless of the database structure. Experiment B favored the network structure, in that the tasks to be performed were relational in nature (that is, retrieval tasks for which one might naturally use a keyword).

Mohageg found that overall, task completion times were significantly faster for subjects using the hierarchical and combination structures than for subjects using either the network or linear structures. However, for those tasks that required subjects to traverse the largest number of links in Experiment B, the network structure allowed significantly faster task completion times.

Generally, uncertainty was lowest for the hierarchical and combination structures, suggesting to Mohageg that "The consistent structure of hierarchical linking allowed for less erratic navigation through the system, whereas the unpredictability of the network structure led to nonuniform navigation patterns." He also noted that subjects using the combination structure typically did not employ the relational links in tasks in which those links would have proved most useful. That is, these subjects seemed unable to recognize when that type of link would have been time-saving.

Mohageg concluded that hypertext databases designed for information retrieval should be hierarchical. He strongly recommended against using relational links in such databases without also including structural links, and he cautioned that when relational links are not suited to tasks, they can degrade users' performance. For example, they can make it more difficult for someone to recognize the optimum path to desired information.

However, Mohageg also found that in terms of task completion time only, the combination structure proved best. Other authors have often cited his paper as evidence that the hierarchical structure is best for informational hypertext documents or web sites. However, because task completion times were fastest for users of the combination structure, Mohageg's results can also be interpreted as evidence favoring this structure—especially considering that information retrieval is unlikely to be the only goal web users have; it is reasonable to expect that many web users will wish to explore a site.

Smith et al. (1997) and Nielsen (1989), after reviewing the hypertext usability literature, concluded that because of inconsistencies among the navigation styles of users, no single optimal structure can be identified. At least one study supports this conclusion. Ellis et al. (1994) asked 25 Information Science graduate students to create links between nodes "whose contents or meaning they decided were related" in five technical documents displayed in the hypertext system Guide (the five documents ranged in size from 23 to 347 paragraphs). The authors observed very little similarity among the links created by the different test subjects, and noted that these results were consistent with previous studies that have found little consistency among different people indexing the same documents. In those previous studies, variation resulted from (1) differences among individual indexers in how they perceived the concepts included in the documents, and (2) differences among indexers in the terms they chose to represent concepts. These results suggest that regardless of how any particular web site is designed, its individual users are likely to differ widely in the degree to which they find the organization of information within the site to be intuitive to them.

Meaningful Organization Hollands and Merickle (1987) obtained results suggesting that the best way to structure a hypertext database or web site may be to concentrate on ensuring that the information it contains is organized in a way that is meaningful to its particular intended audience (although Ellis et al.'s (1994) results suggest that this goal may be difficult to achieve). They asked psychology professors to sort 120 psychology terms into meaningful categories, and then used multidimensional scaling to identify the single categorization of the set of terms most meaningful to the subjects. The subjects identified five top-level categories, which the authors then used to form the top level of a menu hierarchy. The 120 terms were used as options in the second level of the hierarchy. Hollands and Merickle then created two more menu hierarchies with identical structures:

one was alphabetical, in which the top five categories were ranges within the alphabet (e.g., “A to C”); the other was random, with the top five categories given meaningless names (e.g., “Group 1”). The second levels of both of these hierarchies contained the 120 psychology terms. Hollands and Merickle considered the alphabetical arrangement to be a less-natural categorization, and the random arrangement to be a control. Subjects ranging in domain expertise then completed two kinds of tasks using the three hierarchies: (1) term-matching, in which they were given exact terms to locate in the hierarchies, and (2) definition-matching, in which they were given a definition of the term to find. Hollands and Merickle found that for term-matching, subjects using the alphabetical hierarchy were fastest, but that the meaningfully-categorized hierarchy proved fastest for definition-matching. For term-matching, the most expert subjects were only slightly slower when they used the categorized rather than the alphabetical hierarchy. These results suggested to Mayhew (1992) that organizing items in a way that is natural for users improves novices’ ability to find information, and does not detract from experts’ ability to find information. These results also suggest that web sites in which items have been arranged naturally for users will allow people to find information faster.

Card-sorting may be an effective procedure for achieving such a natural arrangement. The purpose of this design technique is to learn how members of a web site’s intended audience would organize the information included in the site; this is assumed to be an organization they would find intuitive. To perform card-sorting, identify a sample of volunteers representative of the site’s intended audience. Present each person with a stack of index cards that you have prepared beforehand, in which each card contains a short phrase describing a potential site component. Ask each interviewee to sort these cards into two groups containing (1) features that should be included in the site and (2) features that should be left out. Also encourage interviewees to create additional cards describing other features they recommend including in the site (keep these cards in the stack for use during later interviews). Next, ask each person to organize the features to be included in the site into groups that seem natural, and then to name each group. Next, ask each person to organize the groups into larger piles, and then to name each pile. For each interviewee, record the features placed within each group and the groups placed in each pile, as well as the names given to the groups and piles.

The result of this set of interviews is a set of two-level, hierarchical “mental models” of the site, one from each interviewee. Especially when there is strong consensus among interviewees, you may be able to discover a natural information organization by simple visual inspection of the card arrangements. Otherwise, you can use a statistical analysis computer program to run a cluster analysis on your data (cluster analysis is a multivariate statistical technique for estimating the relative degree of association among items in a data set). To do this, first generate a similarity matrix from your data by (1) allocating one point to each pair of features for each time both features appear in the same pile, and (2) allocating two points to that pair for each time both features appear in the same group. Then, use your statistical analysis program to run a cluster analysis on this similarity matrix.

Choosing the Number of Levels in the Site

Q. How many levels should a web site contain? Should it be relatively broad and less deep, or vice versa? And how important is this issue?

A. Although few if any researchers have studied the effects of the number of levels in a web site on users' ability to find information in it, people have extensively studied how users search for information in hierarchical menu systems, such as large databases. They have generally found that people can complete searches faster and more accurately when the menu systems are broader (include more choices at each level of the hierarchy) and less deep (contain fewer levels). However, most of these studies have been of systems in which the names of menu options were short and the total number of options was relatively few (usually 64). In a study of a much larger system, people were able to search for information faster and more accurately when the system contained more levels, because the system had been designed to effectively guide them in finding information. The results of this study suggest that shallow, broad web sites are not always better—whether or not a site has been organized in a way meaningful to its users is also a key influence on its usability.

What can web designers do?

- Include no more levels than necessary.
- Use other aids to help people to choose the links that will take them to information they want. In particular, (1) whenever space allows, include link abstracts (a short abstract with each link that explains what users will see if they click that link), and (2), to speed searching through extensive link menus, group related links together into categories, title each category, and display those category titles in the menus.

See Also

Site Topology..20

Making Links Predictable 33

Information Overviews 57

Backgrounder

Most web sites are largely hierarchical in design, including a single home page at the top level, additional index pages at the second level (and at lower levels in extensive sites), and content pages, usually at the third and lower levels. Important web design questions are: What is the optimum number of levels within a single web site? and Is the optimum number different for different kinds of sites?

Many writers have warned designers against including too many levels in a web site, arguing that users should not have to navigate down through many levels to find information. The widely-accepted rule of thumb is that a site should have no more than three levels. For example, Hackos and Stevens (1997) commented that "Regardless of structure most experts agree that users should not have to go through more than three jumps to find the information they need." They added that users will become frustrated if presented with more levels of pages acting as tables of contents rather than providing content. Mayhew (1992) recommended that "Generally, the organization of functionality should reflect the most efficient sequence of steps to accomplish the most likely or most frequent user goal." And Shneiderman (1997) commented that in web sites, "...extra layers are more disorienting than longer index pages..."

Surprisingly, given this degree of consensus, no researchers seem to have systematically examined the effect of number of levels in a web site hierarchy on users' performance. However, many people have compared different configurations of hierarchical menu trees in databases and other online systems to order to find the optimum number of levels in those trees. This issue is commonly described as the "depth-breadth tradeoff" since, in order to present the same number of options to users, you must choose between presenting options in more levels, each smaller in breadth (breadth being the number of options per level), or in fewer levels, each greater in breadth. Norman (1991) identified the main issues involved in the depth-breadth tradeoff:

- As the number of options in any one menu increases, so does visual search time: it takes users longer to scan through the menu and identify the desired option.
- However, as depth of the menu system increases, users must make more choices. When menu options are ambiguous or vague, the chance that any one choice is wrong increases, and users have more opportunities to choose a wrong branch in the system.

In some studies of the depth-breath tradeoff, the task examined—searching through successive menu screens of a database to find a particular piece of information—is similar enough to the task of searching through the pages of a web site for information that extrapolating from these studies to infer effects on web users seems reasonable, at least in lieu of direct web research.

Early Studies of Small Systems One of the earliest studies of the depth/breadth issue was conducted by Miller (1981), who compared 24 college students' search times and accuracy in searching successive menus in a database in order to find a target item at the lowest level. In this experiment, target items were single nouns like "snake" or "France." Miller arranged 64 single-word menu options into four different structures:

(1) six levels of menus, each containing two choices.

(2) three levels of menus, each containing four choices.

(3) two levels of menus, each containing eight choices.

(4) a single menu containing all 64 choices.

Subjects searched through successive levels of menus to find target words. For example, to find "France" in the four-choice, three-level tree, a subject needed to choose "Geography," then "Europe," and then "France."

Miller found that search times were shortest for the two-level menu hierarchy, followed by the three-level structure. Accuracy (measured as percent errors) was substantially better for the two-level structure than for all other structures. He concluded that when design constraints permit, two levels with eight options per menu is the best arrangement for a menu hierarchy. However, Mayhew (1992) reviewed this study and cautioned that it is not possible to conclude from it how people would perform when there are many more than eight choices at each level. Also, because Miller factored system response time out of total search times, it is not possible to know what effect this variable would have had on his results.

Snowberry et al. (1983) performed an experiment similar to Miller's, using essentially the same menu trees of 64 options, except that they also included in their comparisons a single-level menu in which the options were grouped into categories within the menu, rather than arranged randomly. They found that

- Search times were fastest when the menu hierarchy contained only one or two levels than when it was deeper.

- Both speed and accuracy improved when options were categorized rather than arranged randomly.
- Percent errors increased from 4 to 34 percent as the number of levels increased from one to six levels.

The menu systems tested in Miller's and Snowberry et al.'s experiments differed from multilevel web sites in an important respect: menu options were always single words, and the task subjects performed was to search through menus to find a target word at the lowest level of the system. In an experiment using a menu system much more like a web site, Kiger (1984) compared people's searching speed, accuracy, and preferences when they searched menu trees with 2, 3, and 6 levels to find target items within a videotex database of general-interest information.[4] At its lowest level, the database contained 64 pages of information. Twenty-two people (11 Bell Laboratories employees and 11 recruited test subjects) were asked to complete tasks like "You are looking to buy a color television. Use the electronic shopping service to find and buy a color TV." Note that these tasks are very similar to typical web tasks.

Kiger measured each subject's total time to complete each task, including system response time. Unlike Miller and Snowberry et al., he required subjects to recover from any errors they made and included this recovery time in total response times. Despite these differences, his results were similar to those obtained in the earlier experiments:

- Subjects searched fastest and most accurately when the menu tree contained eight items in each of two levels, and least accurately when the menu tree contained six levels.
- Subjects reported that they preferred two-level menu trees and found them easier to use than trees containing more levels.

In Kiger's study, system response times (which he did not report) are likely to have been faster than typical Web response times, suggesting that a similar study performed on the Web might show a larger speed differential among sites with different numbers of levels

[4] A videotex system is a heirarchical database, typically accessed over telephone lines, which people can search by choosing menu options or typing in keywords. Online services such as America Online and Compuserve are the most modern examples of videotex systems (however, the term itself now is rarely used).

(since longer system response times could result in longer total task times for sites with more levels). Kiger suggested that user performance and preference "are more sensitive to the depth of the tree structure than its breadth," and concluded that the experiment indicated "both preference and performance advantages for broad, shallow trees." He suggested that "As a general principle, the depth of a tree structure should be minimized by providing broad menus of up to eight or nine items each." However, Kiger also pointed out that menu item names in his experiment were short, containing between one and five words. He suggested that menus containing longer items might need to contain fewer options.

Studies of Larger Systems With the completion of Kiger's study, an open question remained unanswered: Would results be similar if the database were much larger? To answer this question, MacGregor et al. (1986) performed an experiment similar to Kiger's. However, they used a much larger videotex database, containing 900 document pages. Like Kiger's database, this database contained general-interest information. Different versions included menus containing 2, 4, 8, and 16 options. The authors measured the times taken by 24 student subjects to search through this database for the answers to a set of 40 test tasks such as "Find out what you can use instead of ground beef for a vegetarian hamburger." By using the results from this experiment in a predictive model they had previously constructed, they concluded that for large systems in which menu options are longer than a word or two, the optimal number of menu items per level is four to five. Note that to design an extensive hierarchical information-retrieval system—such as a web site—including only four or five options requires adding more levels to the hierarchy than Miller, Kiger, and Snowberry et al. identified as optimal. MacGregor et al.'s results suggest that the optimum number of levels may be different for large and small systems. In the context of the web, this and the earlier experiments, taken together, suggest that designers of smaller web sites may wish to restrict the number of levels in those sites to just two or three, while designers of larger sites would do better to include more levels.

Another research team demonstrated that design considerations beyond choices of number of levels and options per level can substantially affect a menu system's usability. Savage and Habinek (1984) found that increasing the number of levels in a hierarchical menu system can improve users' performance, as long as the system reliably guides users to options that are relevant for their tasks and away from irrelevant options.

In the first stage of Savage and Habinek's two-stage study, a research team created an initial version of a hierarchical menu interface for a data-processing system designed for three different user audiences (programmers, system operators, and workstation operators). This initial system contained three to four levels of menus and up to nine options per menu screen. Savage and Habinek observed subjects using the system to perform data-processing tasks, and identified problems with the system, such as particular menus that confused users and difficulties people had in choosing the correct initial path into the hierarchy.

In the second stage of their experiment, Savage and Habinek used iterative testing with end users to redesign the system to eliminate the problems they had observed. To do this, they broke up large menus into smaller ones, hence adding levels to the system menu hierarchy. As a result of these changes, task completion times were shortened by 61 percent, navigational errors declined by 93 percent, and the percentage of tasks that were never completed decreased by 75 percent. Savage and Habinek's results suggest that designers of extensive web sites may be able to at least partly mitigate any negative effects of adding levels by incorporating features to help users find the pages they want, such as including link abstracts and categorizing links within menus.

Some Guidelines Paap and Roske-Hofstrand (1988), after reviewing the menu search literature, reported that it is not possible to conclude that there is a consistent tradeoff between menu breadth and depth, since most studies have been of hierarchies with no more than 64 options, and since little is known about the effectiveness of deeper hierarchies in protecting users from straying down irrelevant pathways. They recommended restricting to two levels systems that have "a couple" or more of the following problems: (1) general or abstract menu or descriptor names; (2) top-level menus that are more error-prone than those lower in the hierarchy; (3) users who mostly will be navigating down pathways they have not previously explored; (4) long system response times when tasks place a premium on fast responses; (5) backtracking one step at a time as the only means for users to recover from menu choice errors; and (6) users who must access information located in more than one branch of the hierarchical tree in order to complete a task.

Mayhew (1992) also reviewed the literature on the depth-breath tradeoff, and suggested a simple decision key for menu organization. She suggested that designers (1) choose a maximum breadth of 4 to 10 options per level if menu choice items are complex or you can't meaningfully organize options on the screen; (2) choose a breadth of up to 20 if you can group choices meaningfully and items are short and simple, but users are novices; and (3) if system response time is fast and most users are experienced, present single-level menus if screen size allows it.

Designing Navigation Controls

Making Links Predictable

Q. How can I design links to make it as easy as possible for users to predict their targets?

A. Researchers have consistently found that test subjects using hierarchical menu systems can choose more accurately among alternative menu options when they can view information that helps them predict the effects of choosing any given option. Two ways to present this information to users seem to be helpful:

- Adding a descriptor (a short description of what the user will see if she chooses that option) to each menu option improved people's ability to choose options accurately. This result suggests that including similar descriptors with links on web sites can improve users' ability to find information in web sites. On the web, such descriptors are called **link abstracts**.
- Asking a group of users to choose names for menu options proved to be an especially effective way to improve other people's ability to choose the options they wanted. This result suggests that asking web users for help in naming links also may help to make a web site easy to navigate.

What can web designers do?

- Whenever space allows, include a link abstract with each of your links to explain to users what they will see if they traverse that link. Make link abstracts as comprehensive as possible, rather than just including a few examples of what users will see if they choose a particular link.
- Because asking users seems to be an especially effective way to choose option names, use card sorting, participatory design, or other methods that involve users whenever possible in order to have members of your intended audience name links.

- You also may wish to use **link titles** as additional aids for users who are trying to determine whether or not clicking a link will be useful to them (Nielsen 1998a). Link titles are now supported by the latest versions of some browsers (see below for more details). However, do not rely only on link titles, since most users can't yet see them; use link abstracts as well.

See Also

Choosing the Site's Vocabulary..12

Embedding Links in Text..41

Using Image Links 50

Backgrounder

Paap and Roske-Hofstrand (1988) reviewed studies of people searching through menu systems for target information, and concluded that "Most errors in menu-driven systems occur because the meaning of the options is not clear to the user." Several investigators have examined ways to improve this situation by making the meaning of menu options clearer to users. Their findings are relevant to web site designers because people who cannot readily predict the effects of clicking links within a web site may not be able to navigate to the parts of the site that most interest them.

Link Abstracts In most studies of menu option meaningfulness, people were asked to search through a hierarchical menu system—a task analogous to navigating through a typical, hierarchical web site—to find target information at the lowest level of the hierarchy. In most of these studies, associating more information with each menu option improved people's ability to choose correct options.

In several studies, researchers tried adding different kinds of descriptors to menu items. In these studies, the descriptors used were roughly analogous to the **link abstracts** commonly used on web pages (an example is shown in Figure 1, below).

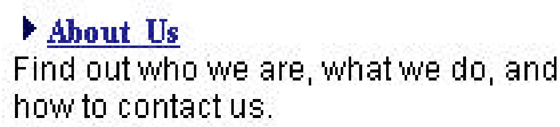


Figure 1. A link abstract displayed on the NOAA HAZMAT home page (response.restoration.noaa.gov).

Lee et al. (1984) investigated the effect of adding a descriptor to each menu option in a videotex system. They worded menu options as keywords or short phrases (using names chosen by an expert panel). In some menus, they added a descriptor to every option name. Each descriptor was a list of the options available at the next-lower level in the hierarchy. Here is an example of one of their menu options followed by a descriptor: "General Interest Guide: News, Weather, Sport, Entertainment, Market Place, Employment, Travel, Leisure, Advice." On other screens, they did not include descriptors with menu options. Test subjects each performed 10 information retrieval tasks using the system. Times to complete each task was measured, and subjects also were asked to indicate their preference for each index page. Generally, they preferred menus with descriptors, and they made 82 percent fewer errors when using menus with descriptors.

In a similar experiment, Snowberry et al. (1985) tested the effect of associating help fields with menu options. In their set of two experiments, 40 college students searched a 64-page hierarchical menu system in order to find target words at the lowest level of the hierarchy. Subjects were able to refer to help fields that contained either the target word, previous selections they had made, or the selections available at the next lower level of the hierarchy. The subjects who were able to see upcoming selections made significantly fewer errors than other subjects. The authors concluded that whenever people must search through multi-level menu systems, the options at each level should contain information about the options at the next level down; they also suggested that including this information is more important when the system contains more levels.

In an experiment intended to identify effective ways to name menu items, Dumais and Landauer (1983) asked more than 700 college students to place 72 items into Yellow Pages categories. Each category was identified either by (1) its name alone, (2) its name and either one or three examples of things in that category, or (3) three examples only. For example, one category in this experiment was "Services Which Come to You"; an example for that category was "Janitor Service." Including three examples as descriptors for each menu item improved subjects' accuracy by only 6 percent. This result suggests that lists of examples may be less effective than other kinds of descriptors.

Would descriptors—that is, link abstracts—be more or less effective on the web? Paap and Roske-Hofstrand (1988) commented that descriptors come with costs: in particular, they are likely to slow the times taken to search each menu, and they take up space. However, on the web, since it takes most users considerable time to download pages, reducing menu selection errors means reducing the number of times the wrong pages are downloaded, and hence might speed overall times to complete information searches in web sites.

What descriptors might be most effective on the web? Paap and Roske-Hofstrand cautioned that adding descriptors that include either a list of options available at the next level or examples may help "either a little or a lot" but that it's not clear just what information should be included in descriptors. However, the evidence from menu system studies seems to be in favor of abstracts that explicitly state what's available at the next level down. The results of one study of web users supports the effectiveness of explicit link abstracts. Spool et al. (1997) observed more than 50 test subjects searching for target information in nine commercial web sites. They reported that sites containing relatively long text links followed by long abstracts rated highest in terms of users' ability to successfully find information they were searching for, and concluded that the most "predictive" type of links were those with relatively long names followed by extensive abstracts. However, they did not quantify their finding.

Another form of link descriptor that has recently been developed for the web, and which is part of the HTML 2.0 specification, is link titles. Nielsen (1998a) recommends using link titles, which are now supported by the latest versions of some browsers, to aid users who are trying to determine whether or not clicking a link will be useful to them. When the HTML code for a link title is included in a page,[5] moving the cursor over the link will cause the link title text to pop up. That text could be, for example, a brief explanation of the item pointed to by the link or an indication of a potential problem, such as "user registration required." Nielsen recommends including only supplementary information in link titles, since users with older browsers cannot view them.

---

[5] Link titles are encoded as follows: <A HREF="/link name/" TITLE="brief description of the link destination">

Naming menu options and links Whether a menu option or web page link will be predictable also depends on its name, and whether that name is meaningful to people. Paap and Roske-Hofstrand (1988) reviewed a series of menu selection experiments and concluded that the cause of errors in choosing among menu options was usually an option name that was not natural or precise. Names that seem obvious to designers may not be obvious to users. In an extensive series of experiments, Furnas et al. (1984) collected evidence to this effect when they asked people such as college students, homemakers, and secretaries to name concepts in different categories (such as text-editing commands and want ad categories). They found that the probability that people would identify the same name for an item ranged from 8 percent for text editing commands to 18 percent for index words for cooking recipes.

An especially effective way to ensure that menu option or link names are meaningful to users is to ask the users themselves. Lee et al. (1984) performed a series of experiments in which they first observed the kinds of errors made by users of hierarchical menu systems, and then measured the effects of changes made to the system in order to reduce the frequency of the observed errors. They observed people using an 79-screen hierarchical menu system that had been designed by experts, and found that 80 percent of all menu choice errors were made on just six menus. When a group of people who were not experienced using this system were asked to rename those six menus, and the experiment was then repeated, using the new menu names, subjects made 40 percent fewer errors.

Dumais and Landauer (1983) obtained a striking result from their study in which students placed items into categories: when they removed the "miscellaneous" option from the list of available categories, the test subjects' accuracy improved by 45 percent. This result suggests that users of menu lists can become confused by an ambiguous name for just one option; the authors concluded that one ambiguous choice can degrade people's ability to understand and use other available choices.

Paap and Roske-Hofstrand (1988) suggested that the potential shortcomings of menu option names fall into two categories: names are (1) "too narrow" when they imply fewer options than are actually available, and (2) "too imprecise" if they imply "more actions or objects than the option controls." They recommended using names that are as precise as possible, and that allow users to see exactly what actions or objects can be accessed by choosing each option, without leaving any out or suggesting things that are not in fact available. However, they note that "This is easier said than done."

Choosing Link Colors

Q. Should designers stick with the default colors for visited and unvisited text links?

A. While some web experts recommend using the default colors as a navigational standard, only a weakly positive effect of keeping default colors has been observed. On the other hand, remaining consistent with the default colors may result in less legible links because people generally, and especially older adults, may have several kinds of difficulties reading saturated, or "pure," blue text (saturated blue is the default color for unvisited web links).

What can web designers do?

- Designers should use their own judgment on this issue. Designers of web sites expected to have many older people in the core audience may especially wish to consider alternative link colors.

See Also

Background and Text Color...64

Backgrounder

The default colors for web page links are blue for non-visited links and purple for visited links. Some authors recommend using these default link colors so that users need not learn the meaning of other link color schemes when they move among sites. For example, Nielsen (1996) identified the use of non-standard link colors as one of the top 10 mistakes in web design, and recommended "Don't mess with these [default] colors since the ability to understand what links have been followed is one of the few navigational aids that is standard in most web browsers. Consistency is key to teaching users what the link colors mean."

Results from one study of web users appear to offer weak support for this position. Spool et al. (1997) observed more than 50 people searching for target information in nine commercial web sites, and reported that using the default link colors had a weakly positive effect on users' abilities to find information. However, they did not quantify this finding.

Legibility Problems Results of studies of human perception, though, suggest that at least some people may have difficulty with the saturated[6], or "pure" blue hue that is the standard default color for unvisited links.

First, it can be difficult to focus precisely on saturated blue objects. Research in visual perception has shown (Silverstein 1987) that

- The normal human eye "focuses blue images in front of the retina, and accommodative adjustments may not be sufficient to bring blue images into clear focus."
- With aging, the eye's range of accommodation becomes smaller, so that older people may have particular difficulty clearly focusing on saturated blue objects.

Saturated blue may be an especially inappropriate color for emphasizing objects such as links: because the eye focuses on blue objects in front of the retina, those objects generally appear to be more distant—and hence less prominent—than objects at the same distance shown in other colors (Murch 1984).

Second, Murch (1984, 1985) reported that the lens and retina absorb much more light from the blue part of the color spectrum than from elsewhere in the spectrum, making the human eye relatively insensitive to blue. This can cause saturated blue objects to look fuzzy, especially in dim light. Lens yellowing further reduces people's ability to discriminate the details of saturated blue objects as they grow older (Murch 1984, 1985).

The results of studies of people viewing saturated blue objects on computer monitors are in accord with these findings from perception research. Silverstein (1987) reviewed the results of two studies of people asked to identify small symbols displayed online in different colors, and reported that people were least able to correctly identify blue symbols. This result suggests that people are also likely to have more difficulty distinguishing saturated blue characters than characters displayed in other colors.

Because of these difficulties, Silverstein (1987) recommended against displaying small, saturated blue objects on computer monitors, while Murch (1984, 1985) and Horton

[6] A saturated, or "pure," hue is one in which the fewest different wavelengths of light are mixed together to produce the color.

(1991) similarly recommended avoiding the use of saturated blue for text, thin lines, or small shapes.

The variety of difficulties people may experience when viewing saturated blue text suggests that saturated blue may not be a good color choice for unvisited links, especially for web sites with core audiences that include older people. Perceptual research suggests that one way web designers can avoid the difficulties with medium blue as a link color is by using dark blue as a link color instead. Adding black to a blue hue “desaturates” the hue; the eye does not experience the same problems viewing text or objects shown in near-black hues that it does when viewing saturated blue text or objects.

Embedding Links in Text

Q. Should links be embedded in text, or kept separate from the text in a menu or navigation bar?

A. Some commentators have speculated that links embedded in the content text on web pages may hinder or distract readers by breaking up the "flow" of text or making it too easy to move to another page. Others have wondered whether embedded links might be too easily missed by readers searching for particular information on web pages.

However, in three studies of people using hypertext documents in which links were either (a) embedded in content text or (b) included in menus separated from the text, researchers found that people using embedded links could navigate through documents faster and more quickly find information they were searching for. Although these experiments were performed on closed hypertext systems, not in the much more open-ended web environment, these results suggest that embedded links do not cause disadvantages for readers.

What can web designers do?

- Assume that there is no need to avoid embedding links within the content text of web pages.

See Also

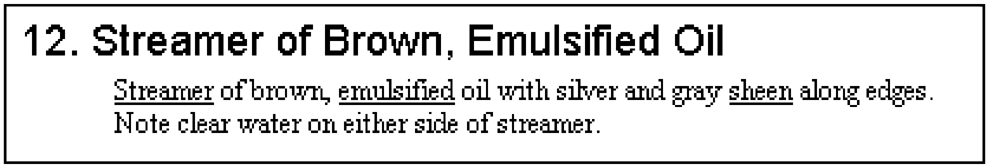
Making Links Predictable 33

Number of Links 54

Arranging Links 45

Backgrounder

Koved and Shneiderman (1986) differentiated between **explicit menus**, usually a list of items from which users select, and **embedded menus**, "where menu items are embedded within the information being displayed on the screen..."


12. Streamer of Brown, Emulsified Oil

Streamer of brown, emulsified oil with silver and gray sheen along edges. Note clear water on either side of streamer.

Figure 2. In this photo caption on a page in the NOAA HAZMAT web site (response.restoration.noaa.gov), "Streamer," "emulsified," and "sheen" are examples of embedded links. Clicking any of these links displays that term's definition.

Koved and Shneiderman pointed out the weaknesses of explicit menus: presenting lists requires space on computer screens, and "extracting information from the original context to construct menu items may mean that the items have to be excessively verbose to be meaningful." But they also suggested possible drawbacks of embedded menus:

- Underlining or highlighting linked text may reduce reading speed and comprehension.
- Embedding links in text may encourage users to explore details of a topic without first developing an overview.
- Embedded menus also may make it easier for users to lose their way within a hypertext document.

Koved and Shneiderman cited two experiments in which embedded and explicit menus were compared. In the first, students explored a hypertext database of information about the University of Maryland Student Union to find answers to a set of 20 questions. Students using embedded menus viewed significantly fewer screens to find answers, answered significantly more questions correctly, and reported preferring embedded to explicit menus. In another experiment, subjects used embedded menus and "pruning" (a system feature allowing users to trim text irrelevant to the task at hand) to search an online hypertext document. These subjects were able to solve a set of problems significantly more quickly while viewing significantly fewer pages than others (presumably, the control group in this experiment used explicit menus, but this was not specifically stated).

In an unpublished study described by Norman (1991), Powell measured the accuracy of people searching through hierarchical databases to find the answers to 20 questions; test subjects each were given 15 minutes to answer as many questions as possible. One group of subjects used embedded menus, and another used vertically-arranged explicit menus. Both groups chose from the same menu options. The group using the embedded menus answered significantly more questions correctly.

Vora et al. (1994) compared four different versions of a 90-node hypertext document on nutrition, including

- Two versions of a graphical interface. Each included a schematic diagram, or "browser graph," showing the hierarchical relationships among available nodes. In one of the two versions, labels indicated the relationship between nodes in a pair (for example, the higher-level node in one pair was titled "Vitamin A," and the label "is one of the" appeared on a line connecting the Vitamin A node to a lower-level node titled "Fat-Soluble Vitamins." The other version did not include these connecting labels.
- Two versions of a textual interface, one with all links embedded in content text, and the other with all links in a horizontal list at the bottom of each screen.

The versions were otherwise identical.

In two test sessions, 40 graduate and undergraduate students each searched through one of the versions of the hypertext to find the answers to 18 questions. Time spent per question and time spent per node were the dependent variables measured. During the first test session, subjects using the labeled graphical version completed searches significantly faster than the other three groups of subjects, and spent less time per node. During the second test session, subjects using the labeled graphical version and subjects using the embedded links version completed their searches faster and spent less time per node than subjects in the other two groups. Subjects also were asked to rate the "user-friendliness" of the version they had used; these ratings did not differ significantly.

The results of this study suggest that at least within a single hypertext document, using embedded links rather than links separated from content text does not cause problems for readers and may in fact be a more effective way to present links. However, an important question remains unanswered: The subject of this experiment was a single, hierarchically-structured document containing no links to other documents. Links within web sites often lead either to different documents within the site or to documents at other sites. It is not clear whether a comparison of different versions of a web site implementing different kinds of links would produce the same results.

In marked contrast to the preceding four studies, Spool et al. (1997) observed more than 50 people searching for target information in nine commercial web sites, and commented that test subjects were less likely to find information they were searching for if access to that information was by embedded links. However, this conclusion does not seem to have been drawn from systematic comparisons.

Arranging Links

Q. How should link menus be arranged on web pages? For example, should they be arranged horizontally across the top or bottom of the page, or as a vertical list along one side of the page?

A. There are two main ways to arrange links on web pages:

- embedded within the text.
- separated from the content text in menus or navigation bars.

The following discussion covers the arrangement of links that are separated from text.

Studies have consistently shown that people can search for information faster when links or menu options are arranged vertically rather than horizontally. This result suggests that including a vertical list of links along the side of each web page to act as a menu of destinations would be helpful to web site visitors.

Path links, which indicate the current level in the site hierarchy and allow users to quickly "jump" to upper levels, represent another kind of link arrangement, which offers theoretical advantages to people searching for information in web sites. No study has yet demonstrated whether this kind of link arrangement is more or less helpful to users than vertically-arranged links. However, one study showed that people can most quickly search for information in a web site when they have more kinds of links to choose from. This result suggests that including not only vertically-arranged links but also path links on pages may aid users.

What can web designers do?

- Offer a vertically-arranged list of links along the left-hand side of web pages.
- Consider also including path links along the tops of pages.

See Also

Embedding Links in Text..41

Number of Links 54

Backgrounder

An important web page design question is whether links should be arranged horizontally along the top and/or bottom of web pages or vertically along the side of pages. (Figures 3 and 4 show examples of the two alternative arrangements.)

Figure 3. A vertically-oriented link menu on a page in the NOAA HAZMAT web site (response.restoration.noaa.gov).

Figure 4. A horizontally-oriented link menu on a page in the NOAA HAZMAT web site (response.restoration.noaa.gov).

Comparing Vertical and Horizontal Arrangements Three groups of researchers observing people searching for target items in menus found that their test subjects could search vertically-arranged menus faster than horizontally-arranged menus. This result suggests that when all else is equal, vertically-arranged links would be most helpful to web site users. Backs et al. (1987) measured response time and accuracy of eight McDonnell Douglas engineers searching horizontal and vertical menus to find a target item. Menus contained either four, 8, or 12 items. Subjects were first presented with a target item, then with a menu display. They pressed a button to indicate when they had found the target item. Subjects located the target items significantly faster when menus were vertically oriented, and their response times were substantially shorter even when menus contained only four items (number of items in a menu and menu orientation did not interact significantly). The difference in response times averaged 150 milliseconds across the three menu sizes.

In a 1986 study described by Tullis (1988), Wolf performed a similar study of people using menu systems. She found that test subjects' search times averaged 37 percent less when menus were organized as vertical, alphabetical lists rather than horizontal lists in running text.

Nygren and Allard (1996) investigated page design for corporate intranets. As part of this study, they asked test subjects to scan horizontally and vertically arranged menus of links on web pages, and then report either the presence or absence of target items (in this written description of a conference presentation, they did not report the number of subjects they tested). Subjects first practiced this task until their search times showed no further improvement, and then were tested. The time to find target items was the dependent variable measured. Nygren and Allard estimated scanning rates (times to find an item) for "close", equidistant, vertically-arranged items to be about 100 ms; for items farther apart (they did not report how much farther apart), scanning times increased to an estimated 220 milliseconds. They also found that the scanning rate for horizontally-arranged items was 1.2 times longer than the scanning rate for vertically-arranged items. From these results, Nygren and Allard concluded that links should be arranged vertically and grouped closely together. However, because this study was conducted with practiced users, it's not possible to conclude that users of non-intranet web sites, who do not typically practice navigating within the same web sites, would respond identically.

Path Links What is not clear from studies comparing vertical and horizontal arrangements is whether **path links** are more or less effective than other link arrangements. Path links are horizontally arranged, with a link to the topmost level of the site usually in the left-most position and links to successively lower levels arranged in order to the right. Typically, the name of the current page occupies the rightmost position in the list. Figure 5, below, shows an example set of path links.

Figure 5. A set of path links on a page in the NOAA HAZMAT web site (response.restoration.noaa.gov).

Unlike other arrangements of links, path links allow users to see where they are within the hierarchy of a web site (assuming that users can readily interpret the arrangement of the links). Path links also allow users to quickly return to higher levels in the site. Countering these potential advantages is the horizontal arrangement of path links, which is likely to slow users down.

Studies to date do not clearly show whether path links are more or less helpful to users than other link arrangements, but results from one study suggest that they may be helpful when offered along with other link arrangements. Bachiochi et al. (1997) compared the effectiveness of three different kinds of navigation aids:

(1) "structure" buttons or links, including "Contents," "Map," "Index," and "Home,"

(2) "local navigation" links or buttons, including "Up," "Previous," and "Next," and

(3) path links or buttons.

In the first phase of this study, 25 non-technical test subjects searched for target information in a 48-page website. Dependent variables measured included the time taken and number of page changes required to find information (the number of errors was apparently not measured). No significant differences were observed among groups of test subjects using different combinations of the three kinds of aids (ranging from having only the local navigation aids and Home link available to having all three kinds of aids available).

In a second phase of the study, 15 subjects (including eight with no previous experience using the Web) searched for information within the same 48-page web site. In this test, browser navigation controls were available to all groups of subjects. In addition, one group was able to use a Home link, another was able to use structure, navigation, and path links, and a third was able to use structure and navigation links fixed at the top of each page (these links were placed within a frame). Generally, the more kinds of navigation controls available, the shorter was the time taken to find information, and the fewer were the page changes made. Because of confounding among types of navigation aids tested, it was not clear from this study whether one or another kind of link (path, structure, or local) is more effective than other kinds, although the authors concluded that local navigation links may confuse users more than path or structure links.

A result from another study suggests that if path links are helpful, they may aid relatively few users. Spool et al. (1997) observed more than 50 people searching for target information in nine commercial web sites. Path links had been implemented on the pages of one of these sites (cnet.com). However, these researchers observed that only one test subject “gave any indication” of noticing these links. This result suggests that path links may be helpful to relatively few users. It is possible that they would prove most helpful to return users of a site, who might eventually notice navigation controls overlooked by one-time visitors. It is also possible that they would be noticed and used by more people if they were to become more common on the web.

When path links are used, should they be placed at the top or bottom of pages? Bachiochi et al. observed that test subjects frequently expressed a preference for navigation controls fixed at the tops of all pages. However, they did not compare users’ preferences for links fixed along the side of the page to links fixed at the top of the page. It is possible that users’ real preference is for navigation controls that, one way or another, never scroll out of view.

Using Image Links

Q. How can I effectively implement icons and imagemaps as links? Or is it better to just stick with text links? What are the pitfalls of image links?

A. To answer these questions, we first must explore two issues: (1) how people interpret graphical information, and (2) potential technical problems with image links.

First, people generally interpret information presented graphically much less consistently than they interpret text. Because of this tendency, it can be very difficult to create icons to represent web page links that most users will understand. It can be especially difficult to create icons meaningful to all visitors to a web site if that site receives many international visitors, since people from different cultures may interpret symbols differently.

Second, there are technical difficulties associated with image links:

- They cannot be traversed by people with images turned off or by sight-impaired people using screen readers.
- The fact that image links do not change color once they have been traversed can cause difficulties for web users who use link color changes as navigation aids.
- Users may have difficulty when the entire area covered by an image link or imagemap is not "clickable."

What can web designers do?

- Because people may not interpret graphic symbols as designers intend, include text labels with all icon and image links.
- When including image maps as navigation aids, either also include a text-only navigation bar containing the same links, or use client-side image maps in which you incorporate ALT tags in the link descriptions for all AREAs of the imagemap. These tags can be read by both text browsers and screen readers (Waters 1997).
- Ensure that the entire area covered by an image map is "clickable," and that clickable areas within the same imagemap are discernible from each other and from other items on the page.

See Also

Making Links Predictable 33

Choosing the Site's Vocabulary...12

Backgrounder

Web site designers often use icons and imagemaps as graphically-pleasing alternatives to text links. However, there are potential problems with replacing text links with image links. In particular, ensuring that all users will interpret a graphic or icon in the same way is more difficult than ensuring that everyone will interpret the same word or phrase similarly,[7] and large graphics such as imagemaps can slow page downloading.

Icon Intuitiveness Rogers (1986) performed a matching study to evaluate the effectiveness of icons intended to represent word-processing actions such as "delete a block of text" and "quit." She asked 60 student subjects to match icons with the commands they were intended to represent. She found that the more abstract or complex the concept represented by the icon, the less likely were test subjects to recognize what it represented. Icons representing concrete objects that were being directly acted on were most often recognized correctly. These results suggest that icons used as direct manipulation controls may be most effective when they represent simple, concrete functions.

Nielsen and Sano (1994) performed an icon intuitiveness study in the course of designing the interface for the Sun Microsystems intranet. They showed icons designed to represent 15 different categories of information to four study participants, and asked the participants to explain what they thought the icons symbolized. Nielsen and Sano found that "people got the general idea" of what some icons represented, but ranged widely in their interpretations of other icons. For example, a toolbox intended to represent a "Specialized Tools" category was interpreted as a toolbox by just one participant and as a briefcase by others, and a set of computer components meant to represent an online product catalog was interpreted as CD-ROM information by some participants, but also

---

[7] Choosing text that people will interpret similarly also can be difficult: see "Choosing the Site's Vocabulary."

as “system-oriented” or “disk” information by others. Nielsen and Sano’s results suggest that because people differ in their interpretations of icons, using only unlabeled icons as links or to represent options is very likely to confuse at least some visitors to a web site.

Labeling Icons Once users have learned the meaning of icons, those meanings may be better remembered than the names of the commands they represent (Horton 1991). Perhaps for this reason, many web and software designers have chosen redundant representation—that is, graphic icons with text labels—in which the text label is intended to “disambiguate” the icon. The disadvantage is that icons with labels take up more space on the screen or web page. The memorability advantage of icons would be important for web sites only when designers expect many people to become frequent visitors; in such cases, there is a usability advantage to constructing memorable interfaces. Otherwise, attempting to design web pages that are easy to understand when first viewed is likely to have a much greater effect on site usability.

Cultural Differences An issue that was not addressed in these studies, probably because they were conducted on software that was not expected to be widely distributed, is the degree to which people from different cultures may interpret icons differently. Waters (1997) pointed out that the blue, fourlegged mailbox that would be easily recognized by U.S. readers would not necessarily be recognized as a mailbox by readers in the U.K., where typical mailboxes are red, cylindrical, and have no legs, or by readers in Japan, where mailboxes are cylindrical and have rounded tops. More research on the meaningfulness of symbols across cultures would be helpful to the web design community.

Technical Issues Besides the question of how images may be interpreted by users, there are at least three potential technical problems involved in implementing image links on web pages. First, image links cannot be traversed by people using text-only browsers or browsers with images turned off, or by people using screen readers. In the case of imagemaps, Waters recommends including either an alternative set of text links or creating a client-side imagemap, and encoding an ALT tag in the link description for each AREA included in the image map. These ALT tags can be read by both text browsers and screen readers.

Second, the results of one study suggest that when image links are implemented on web pages, it is important to ensure either that the entire image is “clickable” or that users can readily distinguish clickable parts of the image. Nielsen and Lyngbaek (1990) performed a field study to investigate the usability of a children’s story presented in HyperCard, and found that children missed many clicks because link anchors did not cover the entire area that the children apparently expected to be the logical or visual anchor point. Waters (1997) made the related recommendation that clickable areas within the same imagemap should be discernible from each other, and Sano (1996) also warned, without presenting evidence, that users can become confused if there isn't a clear visual distinction between controls, such as image maps and buttons, and static visual elements, such as logos and graphics. Nielsen and Lyngbaek's findings lend some support to these recommendations.

Third, Spool et al. (1997), in a study of more than 50 people searching for target information in nine commercial web sites, observed that the fact that image links don't change color once you have traversed them caused problems for users in at least one instance. Likewise, during a usability test of the NOAA HAZMAT web site conducted on my computer, one participant complained that he found it difficult to navigate within the site because he could not use link colors to identify the parts of the site he had visited (nearly all the links had been changed to their "visited" color because I had previously traversed them).

Number of Links

Q. How many links should be on a page? How many is too many?

A. Some researchers have suggested that offering users too many choices at once may confuse them and make errors more likely. If this is so, then presenting more than a certain number of links on each web page would make it more difficult for people to search through the site. But how many links is too many?

In two early studies, people searching for items in hierarchical menu systems found those items faster when the menus were shorter. However, in recent studies of people searching for information in web pages,

- People found information fastest when pages within the site contained more rather than fewer links (these also were the largest pages).
- Presenting alternative links on the same web pages allowed users to complete information searches faster.

What can web designers do?

- In the absence of additional findings, assume that adding more links to pages will not necessarily degrade user's performance.
- Because research results in this area are ambiguous, examine the effects of link number and density in usability tests.

See Also

Arranging Links 45

Embedding Links in Text..41

Page Density 60

Choosing the Number of Levels in the Site 26

Backgrounder

How many controls—text and icon links and image maps—should designers include on each page within a web site? How densely should those controls be arranged on pages?

Perlman (1984) examined the effect of menu list length on search times. He observed subjects searching for target words in menus in which single-word options were randomly ordered, and found that search times increased linearly with menu length. However, alphabetizing options within menus significantly shortened search times.

MacGregor et al. (1986) observed 24 test subjects searching for information in four different versions of a 900-page videotex database. The different versions included menus containing 2, 4, 8, and 16 options. The authors measured the times taken by subjects to search through this database for the answers to a set of 40 test tasks such as "Find out what you can use instead of ground beef for a vegetarian hamburger." By using the results from this experiment in a predictive model they had previously constructed, they concluded that for large systems in which menu options are longer than a word or two, the optimal number of menu items per level is four to five.

Perlman's and MacGregor et al.'s results suggest that when fewer links are displayed on a page, users will be faster at choosing correct links. However, these findings have to be balanced against the fact that on web sites, reducing the number of links per page, especially on index pages, will mean that users must traverse more pages to find particular target items than they would if there were fewer pages, each containing more links, to traverse.

DiPierro et al. (1997) measured the time taken by 12 University of Maryland students to navigate through three sets of web pages by reading and following links, using three different screen sizes. The most links were visible at any given time on the largest screen. They found that students followed links significantly faster when using the largest screen size, and also found that 75 percent of the test subjects preferred the largest window size. They suggested that, based on these results, links on "index" and "table of contents" web pages should be arranged as densely as possible, since it is these pages that users must search to find links of interest.

However, DiPierro et al. examined only the effect of the number of links on the page, not density of links. What we can't know from their results is whether including more links on pages while holding page sizes constant would reduce search times. However, their results suggest that increasing the number and hence the density of links on pages may not necessarily cause problems for users.

Bachiochi et al. (1997) found that presenting more kinds of navigation controls to test subjects reduced times required to find target information. They compared the effectiveness of different kinds of links and buttons allowing users to navigate from page to page and among levels within a 48-page test web site. Fifteen subjects (including eight with no previous experience using the Web) searched for information within the site. Generally, when more kinds of navigation controls were available, subjects took less time to find information and made fewer page changes. This result suggests that including more links on pages does not necessarily impede users' navigation.

In another study, researchers obtained results different from those of DiPierro et al. and Bachiochi et al. Spool et al. (1997) observed more than 50 people searching for target information in nine commercial web sites, and commented that when there were more links on a page, users were less likely to find the information they were searching for. However, they did not quantify this finding; it appears to be an impression rather than the result of systematic study.

Information Overviews

Q. Are site maps, site indexes, or other information overviews helpful to web users?

A. There is some evidence that people performing different kinds of tasks prefer different kinds of overviews. In studies of hypertext systems, people searching for specific information have preferred to use an index-style overview, while people exploring a system preferred either a graphical overview or a guided tour.

In studies of the effectiveness of overviews used on web sites,

- Web users expressed a preference for having an overview available to them, and in a study of people searching for information within a web site, people who chose to use a site map were helped by it.
- However, an analysis of one site server log showed that the site index for that web site, at least, is little-visited relative to other pages in the site.

What can web designers do?

- Designers of informational web sites, in which users are especially likely to search for specific information, may wish to provide site indexes.

See Also

Site Topology...20

Choosing the Number of Levels in the Site 26

Backgrounder

The commonest kinds of information overviews in computer systems are (1) system indexes, which are textual in nature, and (2) system maps, which are graphical. An example of a system index for a web site is shown in Figure 6, below. Shneiderman (1997) noted that a wide variety of other kinds of overviews are possible, though, and listed as examples starfields (zoomable scattergrams of color points), tree diagrams, tree maps (nested rectangles that indicate site hierarchies), parallel coordinates, and network diagrams.

- **About Us**
  - How to Contact Us
  - Our Role in Response
  - Our Current Projects
  - Upcoming Training
- **Aids for Oil Spill Responders**
  - Oil Spill Toolkit
    - Open Water Oil Identification Job Aid
    - The ESI Project
    - Spill Software
  - Publications
    - Response Reports and Case Histories
    - Manuals and Guidelines
    - Reports on Spills and Response Issues
- **Aids for Chemical Accident Responders**
  - CAMEO, ALOHA, and MARPLOT (basic descriptions)
    - ALOHA Technical Description
  - CAMEO Toolkit
    - ALOHA Decision Keys
    - Answers to Your ALOHA Questions

Figure 6. Part of the site index for the NOAA HAZMAT web site (response.restoration.noaa.gov). This is an example of a system index.

Shneiderman commented that empirical testing is needed to "reveal what kinds of overviews are most effective and whether performance times, error rates, or retention are enhanced by certain overviews." Until those studies have been completed, he suggested, web designers can use results from other user interface domains, especially research into menu systems, hypertext, and information retrieval.

Results from two studies of people using hypertext systems suggest that the kind of tasks being performed by users affects the kind of overview they will find most useful, and that people searching for information prefer indexes. Smith et al (1997) reviewed four studies of hypertext usability, and concluded that index navigation aids are preferred by people performing searching tasks, while people exploring the system prefer graphical representations of the information. Hammond and Allinson (as described in Nielsen 1989) found that users exploring a hypertext system in order to learn about a subject made little use of the system's index but heavily used the system's guided tour feature; in contrast, users trying to answer specific questions more often used the index.

One team of researchers found that including overview maps as navigational aids in hypertext documents actually increased users' navigational difficulties. Stanton et al. (1992) observed 24 college students searching through a 42-screen hypertext document to find the answers to a set of questions. Half the students used a version of the document containing an overview map linked to each screen, while the other half used a version without an overview map. These authors found that although the two groups did not differ significantly in their success at finding the answers to the set of questions, the group working without a map traversed significantly more links while completing tasks. Also, subjects were asked to diagram the document structure after completing the task set; the group that had worked without a map produced significantly more detailed and accurate maps. The authors hypothesized that subjects working without a map learned more about the document's structure because they were more actively involved with navigating the document.

Few people have studied whether offering information overviews is helpful to people navigating within web sites, and it also is not yet clear whether particular kinds of overviews are more or less helpful to web users. However, the results of one small-scale study suggest that web users prefer to have an overview of some kind available to them. Nielsen (1997a) performed a series of usability tests of the Sun Microsystems web site, and noted that "Users consistently praised screens that provided overviews of large information spaces."

Results from another study suggest that information overviews aid those web users who choose to use them. Spool et al. (1997) observed more than 50 test subjects searching for target information in nine commercial web sites. One of the sites, Fidelity Investments (www11.fidelity.com:80/), included a site map. Spool et al. reported that "The users who used Fidelity's site map were twice as successful at finding answers on that site as the users who did not." They did not report the percentage of subjects who did and did not choose to use Fidelity's site map.

Some evidence, though, suggests that few web site visitors refer to site overviews. An analysis of the server log for the NOAA HAZMAT web site (response.restoration.noaa.gov) shows that the percentage of web site visitors who make use of that site's index is low. During the 11-week period from March 9 to May 24, 1998, visitors to this web site accessed 48,516 web pages and PDF documents. Of this total, only 1 percent of the accesses were of the site index; the home page was accessed about eight times as often (accounting for almost 8 percent of the total page accesses).

Designing Web Pages

Page Density

Q. On individual web pages, how much of the page should be taken up by white space, and how much by text and graphics?

A. In studies of the effects on users of varying the density of computer screen displays, researchers have generally found that when all else is equal, increasing the amount of text and graphics displayed on a single screen slows people searching for information on that screen. However, when researchers studied people working with the same information displayed either at (1) higher densities on fewer screens or (2) lower densities on more screens, they found that people searched for information fastest when screen densities were high, allowing the same information to be displayed on fewer screens. These findings suggest that web users would be better served by pages that are more densely populated with text and images than is typical of printed documents, as long as those pages are effectively designed.[8]

What can web designers do?

- Use less white space on web pages than is traditional for printed text.
- As suggested by Tullis (1988), minimize the amount of information displayed on pages. He recommended using abbreviations when they are well-known to users, and writing as concisely as possible (Morkes and Nielsen (1998) described how they made a text document more concise in order to present it on the web).

[8] A page is effectively designed if it (1) reveals to its viewers the structure of the information it contains and (2) offers easily-recognized access points into that information. Designers can accomplish these two design goals by techniques such as grouping related information together, using white space to separate groups of related information, and using headings and subheadings to label topic areas.

See Also

Number of Links 54

Choosing the Number of Levels in the Site 26

Writing to be Read Online 16

Backgrounder

Page density is the percentage of a web page that is taken up by characters and graphics; white space is the percentage of the page where nothing is displayed. The two sum to 100 percent.

In early studies of screen density (the density of characters and graphics displayed on a computer screen), researchers consistently found that increasing density slowed the times taken by test subjects to find information on a single screen. For example, Tullis (1988) compared the times taken by test subjects to search for a single item of airline or lodging information on screens arranged in a total of 52 different formats, ranging in density from 10 to 58 percent. Generally, search time increased with density, but the low correlation (0.33) between density and speed suggested to Tullis that other factors interacted with density. He concluded from this experiment that "the tendency for higher densities to be associated with degraded human performance can be mediated by other screen format characteristics."

Guideline authors have proposed screen density thresholds that should not be exceeded, recommending thresholds ranging from 25 to 60 percent. Tullis, after reviewing studies of screen density, concluded that "there is no clear evidence supporting the concept of a specific density 'threshold' (e.g. 25% or 60%) that should not be exceeded," since many other factors—especially arrangement and grouping of items on screens—strongly mediate display effectiveness.

Results of a later experiment suggest that to evaluate the effects of web page density, it is necessary to consider the effects of adding additional pages to a web presentation in order to reduce the density of the pages included in the presentation. Staggers (1993) measured the effects of screen density on the task performance speed, accuracy, and subjective satisfaction of 110 nurses working with alphanumeric patient information presented on computer screens. She operationalized screen density as the number of displayed characters divided by the total number of text spaces available on the screen. She compared low-density (27 to 32 percent), moderate-density (33 and 41 percent), and high-density (58.5 percent) screens. In all cases, test subjects viewed the same information, but while a complete set of patient information was presented on a single high-density screen, it was presented on two moderate-density screens and three low-density screens (information was segmented among moderate- and low-density screens by logical groupings). Information was visible on all screens without scrolling, and system response time was "sub-second."

Staggers found that her subjects' task performance speed was significantly faster when they used high-density screens, and that performance speed also was significantly faster on moderate-density screens than on low-density screens. Accuracy and subjective satisfaction were almost unaffected by screen density.

In Staggers' experiment, test subjects initially practiced at the tasks they were then tested on, and then repeated task sets over a total of five test trials. Hence, they were much more practiced than typical web page readers, who are likely to visit only a few web sites frequently. However, Staggers' results held for initial trials as well as later repetitions, suggesting that the effects of screen density are similar for novice and practiced users. Her results suggest that, at least up to a point, presenting information more densely on fewer pages can allow web users to find information more quickly, without risking the loss of either search accuracy or subjective satisfaction. Because she did not include scrolling as an independent variable in her study, and because system response time in her study was much faster than response times typical of the web, it seems reasonable to expect that higher densities may be even more beneficial on the web that in a system like the one she studied.

There appear to have been no systematic studies of the effects of page density on web users' performance. However, Spool et al. (1997) studied more than 50 people searching for information on nine commercial web sites, and reported that "The more white-space there was on a site, the less successful users were at finding information." However, they did not quantify this result. These authors speculated that while white space may aid readers, it may not be helpful to people skimming pages, who will be most successful when they can skim the most material the most quickly.

Tullis (1988) recommended that to minimize screen density as much as possible, designers should use task analyses to ensure that only the information needed by users to perform expected tasks is presented on screens. He suggested several possible design measures to aid in this: (1) using abbreviations when they are needed to save space, are well known to users, and are used consistently (however, an abbreviations dictionary should also be made available); (2) avoiding unnecessary detail (he gave the example of presenting numeric information to as few decimal places as possible); (3) using concise wording; (4) using familiar data formats in order to avoid labeling well-known items such as street address, city, state, and zip code; and (5) using tabular formats for data when possible, since this allows column headings to be used and hence eliminates the need to label each data element presented.

Background and Text Color

Q. What are the best colors to use for page backgrounds and text? Are there background and text color combinations that should be avoided? And should designers choose dark text on light backgrounds or vice versa?

A. Researchers studying human vision and people reading online texts have obtained some key findings useful to web site designers:

- The most effective color combinations are those in which text and background colors contrast most strongly in brightness as well as hue. Text is as legible as possible when presented either as white characters on a black background or as black characters on a white background.
- Blue and cyan, the shortest-wavelength colors, are poor choices for text color (unless these colors are darkened by adding black to them), but good as background colors.
- While people reading light characters on dark backgrounds for long periods reported less visual fatigue than people reading dark characters on light backgrounds, people scanning online text for symbols or words could complete their tasks significantly faster when viewing dark characters on light background.

What can web designers do?

- Since most web users scan pages rather than reading extensively, choose a combination of dark characters on light backgrounds, unless site visitors are likely to spend extensive periods of time reading texts.
- To ensure the legibility of text, choose a combination of black or dark-colored text on a white or very light-colored background.
- Avoid using light or medium ("pure") blue text.[9]

[9] Because blue is the default color for unvisited links, web designers who expect to attract older readers also may wish to consider using different link colors.

See Also

Backgrounder

Research results in two broad areas are relevant to the question of what colors are best for text and page backgrounds of web sites:

- First, some researchers have evaluated the effects of **display polarity** (showing light characters on dark backgrounds or vice versa) on people reading text or searching for information on computer monitors.
- Second, some researchers have evaluated the effectiveness of different combinations of text and background color.

Display Polarity Researchers studying the effects of display polarity on readers have obtained conflicting results. Cushman (1986) observed 60 subjects who each spent 80 minutes reading general-interest articles on VDTs. Half the subjects read light characters on a dark background, and half read dark characters on a light background. Cushman assessed reading comprehension by asking subjects to answer a set of questions about the articles they had read. He also assessed visual fatigue, by asking subjects to report the extent of the visual fatigue they felt as well as their general feelings (on a scale from extreme well-being to extreme tiredness). Cushman found that subjects reading light characters on a dark background reported no more fatigue than subjects in a previous experiment who had read printed versions of the same articles. Subjects reading dark characters on a light background reported significantly more visual fatigue. Reading comprehension was similar for readers of both online and print text. Cushman commented that a possible reason for the greater visual fatigue of readers of dark characters on a light background was that the 60-Hz frame repetition rate of the VDTs used in the experiment was not fast enough to eliminate screen flicker, which is more pronounced in displays with light backgrounds.[10]

[10] The vertical refresh rate of current monitors, which is analogous to frame repetition rate, is substantially faster than the 60 Hz repetition rate in this experiment: the Video Electronics Standards Association has established minimum refresh rates of at least 70 Hz for current monitors.

Snyder et al. (1990) compared the response times and error rates of 54 subjects (1) finding and identifying a single character or symbol on a display containing between 70 and more than 150 other characters or symbols and (2) finding out-of-place words within paragraphs. Subjects viewed either dark characters on a light background or light characters on a dark background. In a series of three experiments, both speed and accuracy were significantly better when subjects viewed dark characters on a light background.

No unambiguous implication for web design emerges from these two studies. Perhaps the best strategy for web designers is to choose between the alternative polarities in a task-oriented way: if people are expected to spend more than a few minutes reading text on the pages of a site, the combination of light text on a dark background may leave them less fatigued. If, in contrast, people are expected to scan the pages of the site relatively briefly and/or perform quick searches for information, choosing the opposite combination of dark text on a light background may be best.

Results of three recent studies of people using web sites suggest that, in fact, most web users briefly scan text rather than reading it for extended periods. Morkes and Nielsen (1997) found that "79 percent of our test users always scanned any new page they came across; only 16 percent read word-by-word." Bachiochi et al. (1997) observed that their test participants "do not read blocks of text; they jump around and react to links." Spool et al. (1997) commented that generally, test participants skimmed pages, looking for relevant information; only when they found an item of interest did they read text completely. These findings suggest that choosing dark characters on light backgrounds may be more helpful to most web users.

Text and Background Color Combinations Because physiological, psychological, and cognitive factors interact in complex ways to influence how people perceive and interpret color, it is not possible to present unambiguous, simple rules for combining colors on web pages or in other presentations. However, it is possible to infer some very general guidelines from studies of human perception and of people viewing different color combinations of online text and screen backgrounds.

Horton (1991), discussing the use of color in online presentations, emphasized that the best way to make text legible is to ensure a lightness contrast ratio between text and background of at least 7 to 1.[11] He also pointed out that the surest way to ensure sufficient contrast is to present either black text on a white background or white text on a black background. Murch (1984) described the physiological basis of this need for high contrast: the eye must detect an edge in order to perceive a shape. “Edges guide the eye’s accommodation mechanism, which brings images into focus on the retina.” He noted that in order for an object to appear in sharp focus on the retina, there must be a brightness difference between the object and its background.[12] When the object and background differ only in color, not in brightness, the eye cannot sharply focus on the edge between them.

To identify the most effective color combinations for computer display, Lalomia and Happ (1987) studied the effectiveness of different combinations of text and background color. They measured the response times taken by 40 test subjects to recognize individual letters on screens with 120 different combinations of text and background color.[13] They also asked test subjects to rate each color combination by answering the question, “How long would you work with an application which utilized the [color combination]?” Lalomia and Happ considered the effectiveness of a particular color combination to be a function of both response time and subjective preference.

Generally, their results supported Horton’s recommendation to ensure sufficient lightness contrast: they found that the poorest color combinations had low contrast between text and background colors (however, some relatively low-contrast combinations did well in this experiment). They also found that nearly half the optimal combinations included either black or white combined either with each other or with a color. The top 20 percent of the 120 color combinations tested in this study are shown in Table 4 below.

---

[11] Lightness is the amount of either white or black in the color; a lighter color has more white in it.

[12] Lightness is a property of objects. Brightness is affected by lightness, and also by the intensity of light illuminating the object; a brighter object also appears to be lighter.

[13] Response time was measured as follows: each subject pressed the space bar as soon as she had identified each displayed letter. The letter was then masked, and the subject typed the letter she believed she had seen.

Table 4. The color combinations rated as most effective (Lalomia and Happ 1987). In each row of the list below, the text color is listed first, followed by the background colors that were effective in combination with that text color.

Black with cyan, magenta, or white
Blue with white
Green with black, blue, or red
Cyan with black
White with blue
Light green with blue
Light cyan with black, blue, cyan, red, or magenta
Light magenta with black or green
Yellow with black, blue, cyan, or magenta
High-intensity white with black, green, or cyan

How should web designers interpret these findings? People visiting web sites are not identifying individual letters, but instead are scanning or reading text. Mayhew (1992) reviewed Lalomia and Happ's study and commented that because recognizing individual letters on screens is a different task from reading and scanning text, "generalizations from these data should be made with caution." However, she suggests using Lalomia and Happ's results as tentative guidelines for color choices.

However, Lalomia and Happ's findings contrast with the findings of other researchers in at least two important ways, complicating generalization from their results. First, although blue, cyan, and light cyan appear as text colors in three of the most effective color combinations shown in Table 4, results of perception research suggests that blue and cyan are not good colors for text. Silverstein (1987) reviewed the results of two studies of people asked to identify small symbols displayed online in different colors, and reported that people were least able to correctly identify blue symbols. This result suggests that people are also likely to have more difficulty distinguishing blue characters than characters displayed in other colors. It accords with the results of studies of human perception, described by Murch (1984, 1985), which indicate that

- Because the short-wavelength colors blue and cyan focus in front of the retina, objects in these colors appear defocused (fuzzy).

- Because blue photopigment is not found in the fovea (the small area at the center of the retina, where vision is most acute), small blue objects can seem to disappear when we try to focus on them.
- Because the human lens absorbs more light energy in the blue region of the visible spectrum than elsewhere in the spectrum, we are less sensitive to blue and cyan than other colors. Also, the lens yellows with age, making older people more insensitive to blue.

Second, although the combination of light cyan with red also appears as one of the most effective color combinations in Table 4, the results of research in human perception suggest that at least under some conditions, this also may not be an effective combination. Because the human eye focuses on short-wavelength colors in front of the retina and on long-wavelength colors behind the retina, Murch (1984, 1985) recommended avoiding combining colors from opposite ends of the color spectrum (termed “spectral extremes”), such as cyan or blue with red or yellow with purple, in order to reduce visual fatigue caused by constant refocusing.

However, because studies intended to measure the strength of this effect have produced contradictory results, it’s not possible to predict how visually fatigued web users may become when viewing text and backgrounds displayed in these combinations, and it seems possible that in practice, the effect generally may be small. Matthews (1987) observed 120 college students proofreading text displayed in different combinations of text and background colors on a CRT monitor. Each test subject worked for about 2.5 hours. Matthews measured the amount of text proofread and the number of errors detected, and also tracked reported discomfort symptoms and asked subjects for subjective ratings of their comfort. Text was displayed in red, yellow, green, blue, or white on blue, red, green, or black backgrounds. Subjects proofreading either red text on a blue background or blue text on a red background (a pairing of spectrally-extreme colors, in both cases) were less accurate and reported more discomfort than subjects working with other color combinations. However, Matthews et al. (1989) observed 105 student subjects scanning a CRT monitor for a target symbol displayed among nontarget symbols and then performing a simple decision task using information associated with the found symbol. Symbols were displayed in different combinations of text and background colors, this time including red on blue (but not blue on red). Although the authors expected people working with this color combination to perform more poorly than others, since it is a combination of spectrally-extreme colors, they found no significant performance difference between these and other subjects.

Text Legibility

Q. What steps can web designers take to ensure that text is legible?

A. Legible text is easy to decipher. Legibility is affected by factors such as the size of individual letterforms, line length, case (uppercase, lowercase, or mixed case), and style (plain, italicized, or bold). While an extensive body of research on the legibility of printed text has been accumulated, the legibility of online text has been little studied, and guidelines for ensuring legibility of online text can't yet be developed. However, clues from a few studies are available to web designers:

- People's reading speed and accuracy improved as font size was enlarged; these people also reported preferring larger font sizes.
- Italicizing online text reduced its legibility.
- Although uppercase text is read more slowly, readers could more quickly find a word on a computer screen when it was in all capitals. This result suggests that presenting headings and subheadings in uppercase may be helpful to readers.
- The optimum line length of online text was found to be 16 to 36 degrees of visual angle, roughly corresponding to a line length range from 26 to perhaps as many as 80 to 132 characters.

What can web designers do?

- Avoid using the FONT SIZE tag to specify relative font sizes of less than 3 for body or display text.
- Avoid specifying font sizes in points when preparing HTML code. Text specified in points appears on a Macintosh to be about three-quarters the size of text at the same point size displayed in Windows.
- Use the BLOCKQUOTE tag or tables to restrict line length.
- In addition, ensure that a strong lightness contrast exists between text and page background (see "Background and Text Color" for more detail).

See Also

Writing to be Read Online 16

Background and Text Color...64

Backgrounder

Font size Factors such as text line length, monitor resolution, the distance of the reader from the monitor screen, screen luminance, and text and background contrast influence how a designer's choice of font size will affect readers. For this reason, it is not possible to firmly conclude which font sizes are best for online text. However, Schriver (1997), citing recommendations of the American National Standards Institute, made a general recommendation that in online presentations, body text should be 10 or 11 points[14] in size and display text 14 points or larger in size. Hackos and Stevens (1997) recommended using a sans serif font at 11 to 12 points for online text. They also suggested using a larger font size if the audience for a product is expected to include many middle-aged and older people. However, these authors did not cite research results.

The results of one study suggest that within the range of font sizes typically chosen for displaying body text in computer applications, larger sizes are better than smaller sizes for extended reading. Tullis et al. (1995) observed 15 volunteers counting the number of typographical errors in paragraphs of text displayed on gray and white backgrounds in a Microsoft Windows application in serif and sans serif fonts, at font sizes ranging from 6 to 9.75 points (defined according to Windows standards). Test subjects' task times and accuracy were measured. Generally, both speed and accuracy increased as font size increased, suggesting that at least within this size range, larger font sizes improve reading performance. The test subjects also expressed a preference for larger font sizes.

Two technical issues affect the size of online text viewed by different people:

- Because people use monitors with different resolutions to view web pages, text of the same point size will look larger or smaller, depending on the resolution at which it is viewed. For web designers, the best design strategy

---

[14] The standard for printed texts is that 72 points make up an inch.

may be to avoid specifying the size of body text, instead allowing it to default to whatever size the user has chosen.

- Font sizes specified in points are interpreted differently on Macintoshes and in Windows. A 9-point Windows font appears to the viewer to be approximately the same size as a 12 point Macintosh font. For this reason, avoid specifying the absolute size of text on your pages in points, since letterforms specified in this way will appear to be different sizes when viewed in the two operating systems.[15]

Case In studies of people reading printed texts, Tinker (1955) and Poulton and Brown (1968) found that mixed upper- and lower-case text is read about 13 percent faster than text presented in all capital letters. Tullis (1988) interpreted this finding to mean that most text displayed on computer screens should be in mixed case. However, Vartabedian (1971) found that the time test subjects took to find a word on a computer screen was about 13 percent shorter when the word was upper-case than when it was lower-case. Vartabedian's result suggests that using uppercase letterforms for headings and subheadings may improve readers' ability to see the structure of an online document by visually emphasizing headings and subheadings.

Style Based on research by McVey (1985), who found that italic lettering is difficult to read when presented in electronic media, Schriver (1997) recommended avoiding using italic text in online presentations.

Line Length Optimum line length depends on font size (point size) and leading. Generally,

- The tighter the leading between lines in a block of text, the shorter should be the length of the lines of text.
- The larger the point size of the text, the longer can be line lengths.

---

[15] For example, when using cascading style sheets, avoid using the font style property to specify the absolute size of text (an example would be <STYLE="font-size: 7pt">); consider leaving font size unspecified or specify relative size or size in percent (for example, <STYLE="font-size: +1"> or <STYLE="font-size: 200%">).

However, it is not yet possible to specify leading in HTML, and, although web designers can specify font size, their choices can be overridden by users. Hence, it may be possible only to specify an optimum line length range for online text.

An optimum line length range was identified in one experiment. Gould and Grischkowsky (1986) measured reading rates for 69-character lines that subtended different visual angles (because they manipulated the size of the characters in these lines). They found that reading speed and accuracy were roughly the same when line length ranged from about 16 to 36 degrees, but fell off at larger and smaller angles. Tullis (1988) interpreted this finding to mean that lines displayed on computer screens should range in length from about 26 to perhaps as long as 80 to 132 characters (the maximum line length on most monitors at that time).

Blinking Text and Animated Graphics

Q. How do blinking text and animated graphics affect readers of web sites?

A. There is a consensus among interface experts that using blinking and animation when they are not needed to convey information can distract or irritate readers. However, early studies suggest that blinking of text may be less detrimental than has been assumed. In one study, people's ability to search for words displayed online was not substantially affected when other screen elements blinked randomly. In a related study, blinking appeared to reduce the legibility of large blocks of text by about 10 percent. It is not clear whether or not blinking text irritated readers.

However, while no study has examined users' subjective preferences for blinking or non-blinking text, one group of researchers observing people searching web pages for information reported that subjects appeared to be irritated by animated graphics. On these pages, the animations were purely decorative, and did not relate to the information that test subjects were searching for. It remains unclear whether animations used to convey information also would irritate users.

What can web designers do?

- Use animated graphics only to convey information.
- Because text encoded with the BLINK tag cannot be read by screen readers, avoid using the BLINK tag.
- In particular, avoid having large blocks of text blink, since this appears to reduce legibility.

See Also

Text Legibility..70

Backgrounder

On many web pages, animated graphics are employed to capture readers' attention. The HTML BLINK tag also is often used to highlight key words or phrases. How effective are these techniques? Do they have unintended consequences?

There is consensus among interface experts that flashing or blinking of screen elements should be used with discretion. For example, Tullis (1988) commented that flashing on computer screens "will almost certainly get the user's attention. It will also almost certainly annoy the user if it is used excessively or cannot be turned off." Likewise, Mayhew (1992) warned designers to "Use blinking with discretion...because it is hard to ignore and can be annoying if it is not really necessary to attend to the item."

In two early studies, researchers studied the effects of blinking text on users searching for words or letters on a CRT display. Smith and Goodwin (1971) observed 10 men searching CRT displays for target 4-letter words, comparing screens on which a random selection of words blinked and screens on which no words blinked. They found that search times "weren't markedly different" for the two types of displays, and concluded that there had been "no marked interference attributable to irrelevant blink." They also reported that search times were about 50 percent faster when either the target word blinked or all words except the target word blinked. This study leaves open some important web design questions. For example, while these test subjects were searching for single words, most web users are searching for larger chunks of information. It is not clear whether they would be similarly unaffected. Also, Smith and Goodwin did not examine subjective preferences, so it is not possible to determine from their study whether people found the irrelevant blinking irritating. However, overall, these results suggest that using blinking text on web pages may be less detrimental to most users than has been assumed.

In a related study, Smith and Goodwin (1972) observed 12 subjects scanning texts displayed on CRTs to find randomly-inserted letter-substitution errors. Each subject scanned a total of 48 text passages averaging 132 words in length, in which 10 letter substitutions had been arranged. Subjects called out substitutions as they detected them. Smith and Goodwin found that response times were about 10 percent slower when the displayed text blinked, compared with non-blinking text. The results of this study suggest that blinking reduces the legibility of long blocks of text.

Although experimental evidence to date does not suggest that displaying blinking words substantially reduces usability, there is an important reason to avoid using the BLINK tag when preparing web pages. Text encoded with this tag cannot be read by screen readers (Waters 1997), creating difficulties for visually impaired web users.

Perhaps only in one study has the effect of animation on web users been observed. Spool et al. (1997) studied more than 50 users searching for information in nine commercial web sites, and commented that users found animations "uniformly irritating." Some test subjects scrolled animations off the browser screen or covered them up with their thumb or hand. However, on the web sites tested in this study, all animations were decorative; none were used to support content. Whether animations used to convey information would add or detract from a site's usability remains an open question.

References

Bachiochi, D., M. Berstene, E. Chouinard, N. Conlan, M. Danchak, T. Furey, C. Neligon, and D. Way. Usability studies and designing navigational aids for the World Wide Web. Computer Networks and ISDN Systems 29: 1489-1496.

Backs, R.W., L.C. Walrath, and G.A. Hancock. 1987. Comparison of horizontal and vertical menu formats. *Pp*. 715-717. *In* Proceedings of the Human Factors Society 31st Annual Meeting, Vol. 1. New York City, October 19-23, 1989. Santa Monica, CA: The Human Factors Society.

Barber, R.E., and H.C. Lucas, Jr. 1983. System response time, operator productivity, and job satisfaction. Communications of the ACM 26(11): 972–986.

Bloom, C. P. 1987-88. Procedures for obtaining and testing user-selected terminologies. Human-Computer Interaction 3: 155-177.

Chapanis, A. 1965. Words, words, words! Human Factors (February 1965): 1-17.

Cushman, W.H. 1986. Reading from microfiche, a VDT, and the printed page: subjective fatigue and performance. Human Factors 28(1): 63-73.

DiPierro, C., G. Nachman, and B. Raderman. 1997. Screen size and web browsing. http://www.otal.umd.edu/SHORE/bs03/index.html (January 21, 1998).

Dumais, S.T., and T.K. Landauer. Using examples to describe categories. *Pp*. 112-115. *In* Human Factors in Computing Systems: the Proceedings of CHI'83. New York: ACM Press.

Ellis, D., J. Furner-Hines, and P. Willett. 1994. On the creation of hypertext links in full-text documents: measurement of inter-linker consistency. Journal of Documentation 50(2): 67–98.

Furnas, G.W. 1997. Effective view navigation. *In* CHI 97 Electronic Publications: Papers, at http://www.acm.org/sigchi/chi97/proceedings/paper/gwf.htm (August 20, 1997).

Furnas, G.W. 1985. Experience with an adaptive indexing scheme. *Pp*. 131-135. *In* L. Borman and B. Curtis (eds.). Human Factors in Computing Systems: CHI '85 Conference Proceedings. New York: ACM Press.

Furnas, G.W., T.K. Landauer, L.M. Gomez, and S.T. Dumais. 1987. The vocabulary problem in human-system communication. Communications of the ACM 30(11): 964-971.

Furnas, G.W., T.K. Landauer, L.M. Gomez, and S.T. Dumais. 1984. Statistical semantics: analysis of the potential performance of keyword information systems. *Pp*. 187-242. *In* J.C. Thomas and M.L. Schneider (eds.). Human Factors in Computing Systems. Norwood, NJ: Ablex.

Gould, J.D., and N. Grischkowsky. 1986. Does visual angle of a line of characters affect reading speed? Human Factors 28(2): 165-173.

GVU Center, Georgia Institute of Technology. 1997. GVU's 8th WWW User Survey. http://www.gvu.gatech.edu/user_surveys/survey-1997-10/ (March 14, 1998).

Hackos, J.T., and D.M. Stevens. 1997. Standards for Online Communication: Publishing Information for the Internet/World Wide Web/Help Systems/Corporate Intranets. New York: John Wiley & Sons, Inc.

Hardman, L. 1989. Evaluating the usability of the Glasgow Online hypertext. Hypermedia 1(1): 34-63.

Hollands, J.G., and P.M. Merickle. 1987. Menu organization and user expertise in information search tasks. Human Factors 29(5): 577-586.

Horton, W. 1991. Illustrating Computer Documentation: The Art of Presenting Information Graphically on Paper and Online. New York: John Wiley & Sons, Inc.

Instone, K. 1998. Usable Web: Guide to Web usability resources. http://usableweb.com/ (April 29, 1998).

Kiger, J.I. 1984. The depth-breadth trade-off in the design of menu-driven user interfaces. International Journal of Man-Machine Studies 20: 201-213.

Koved, L., and B. Shneiderman. 1986. Embedded menus: selecting items in context. Communications of the ACM 29(4): 312-318.

Lalomia, M.J., and A.J. Happ. 1987. The effective use of color for text on the IBM 5153 color display. *Pp*. 1091-1095. *In* Proceedings of the Human Factors Society 31st Annual Meeting.

Lee, E., T. Whalen, S. McEwen, and S. Latremouille. 1984. Optimizing the design of menu pages for information retrieval. Ergonomics 27: 1051-1069.

MacGregor, J., E. Lee, and N. Lam. 1986. Optimizing the structure of database menu indexes: a decision model of menu search. Human Factors 28(4): 387–399.

Matthews, M.L. 1987. The influence of colour on CRT reading performance and subjective comfort under operational conditions. Applied Ergonomics 18(4): 323-328.

Matthews, M.L., J.V. Lovasik, and K. Mertins. 1989. Visual performance and subjective discomfort in prolonged viewing of chromatic displays. Human Factors 31(3): 259-271.

Mayhew, D.J. 1992. Principles and Guidelines in Software User Interface Design. Englewood Cliffs, NJ: Prentice-Hall, Inc.

McVey, G. 1985. Legibility in film-based and television display systems. Technical Communication 32(3): 21-28.

Miller, D.P. 1981. The depth/breadth tradeoff in hierarchical computer menus. Pp. 296 –300. In Proceedings of the Human Factors Society, 25th Annual Meeting. Santa Monica, CA: The Human Factors Society.

Miller, R.B. 1968. Response time in man-computer conversational transactions. Fall Joint Computer Conference–Proceedings: 267–277.

Mohageg, M.F. 1992. The influence of hypertext linking structures on the efficiency of information retrieval. Human Factors 34(3): 351-367.

Morkes, J., and J. Nielsen. 1997. Concise, scannable, and objective: how to write for the Web. http://www.useit.com/papers/webwriting/writing.html (October 6, 1997).

Morkes, J., and J. Nielsen. 1998. Applying writing guidelines to web pages. http://www.useit.com/papers/webwriting/rewriting.html (January 8, 1998).

Murch, G.M. 1985. Colour graphics: blessing or ballyhoo? *Pp*. 333-341. *In* Baecker, R.M., and W.A.S. Buxton (eds.), Readings in Computer-Human Interaction. Los Altos, CA: Morgan Kaufmann Publishers, Inc.

Murch, G.M. 1984. Physiological principles for the effective use of color. IEEE CG&A, November 1984: 49-54.

Nielsen, J. 1998a. Using link titles to help users predict where they are going. http://www.useit.com/alertbox/980111.html (January 13, 1998).

Nielsen, J. 1997a. Report from a 1994 usability study. http://www.useit.com/papers/1994_web_usability_report.html (June 11, 1997).

Nielsen, J. 1997b. How users read on the Web. http://www.useit.com/alertbox/9710a.html (October 2, 1997).

Nielsen, J. 1996. Top ten mistakes of web design. http://www.useit.com/alertbox/9605.html (March 1, 1998).

Nielsen, J. 1995. Multimedia and Hypertext: The Internet and Beyond. Boston: Academic Press.

Nielsen, J. 1993. Usability Engineering. Boston: Academic Press.

Nielsen, J. 1989. The matters that really matter for hypertext usability. Pp. 239–248. In Proceedings of the ACM Hypertext '89 Conference. New York: ACM Press.

Nielsen, J., and U. Lyngbaek. 1990. *Pp*. 64–71. *In* R. McAleese and C. Green (eds.), Hypertext: State of the Art. Norwood, NJ: Ablex Press.

Nielsen, J., and D. Sano. 1994. SunWeb: User Interface Design for Sun Microsystem's Internal Web. http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/HCI/nielsen/sunweb.html (March 1, 1998).

Norman, L.K. 1991. The Psychology of Menu Selection: Designing Cognitive Controls at the Human-Computer Interface. Norwood, NJ: Ablex Press.

Nygren, E., and A. Allard. 1996. Between the clicks: skilled users scanning of pages. Paper presented at Designing for the Web: Empirical Studies. October 30, 1996, Redmond, WA. http://microsoft.com/usability/webconf/nygren.rtf (January 21, 1998).

Papp, K.R., and R.J. Roske-Hofstrand. 1988. Design of menus. *Pp*. 205-235. *In* M. Helander (ed.). Handbook of Human-Computer Interaction. Amsterdam: North-Holland.

Patterson, J.F., and C. Egido. 1987. Video browsing and system response time. *Pp*. 189–198. *In* D. Diaper and R. Winder (eds.). People and Computers III. British Computer Society Workshop Series. Proceedings of the Third Conference of the British Computer Society Human-Computer Interaction Specialist Group, University of Exeter, 7–11 September, 1987.

Perlman, G. (moderator). 1990. Evaluating hypermedia systems (panel). *Pp*. 387–390. *In* Human Factors in Computing Systems: Empowering People, CHI'90 Conference Proceedings. New York: ACM Press.

Perlman, G. 1984. Making the right choice with menus. *Pp*. 317-321. *In* B. Shackel (ed.). Human-Computer Interaction–Interact '84. Proceedings of the IFIP Conference (London, September 4-7, 1984). New York: North-Holland.

Poulton, E.C., and C.H. Brown. 1968. Rate of comprehension of an existing teleprinter output and of possible alternatives. Journal of Applied Psychology 52: 16-21.

Remde, J.R., L.M. Gomez, and T.K. Landauer. 1987. SuperBook: an automatic tool for information exploration-hypertext? *Pp*. 175-188. *In* Hypertext '87 Conference Proceedings, November 13-15, 1987, Chapel Hill, NC.

Rogers, Y. 1986. Evaluating the meaningfulness of icon sets to represent command operations. *Pp*. 586-603. *In* M.D. Harrison and A.F. Monk (eds), People and Computers: Designing for Usability, Proceedings of the Second Conference of the British Computer Society Human Computer Interaction Specialist Group. University of York, 23-26 September, 1986.

Sano, D. 1996. Designing Large-Scale Web Sites, a Visual Design Methodology. New York: John Wiley and Sons, Inc.

Savage, R.E., and J.K. Habinek. 1984. A multilevel menu-driven user interface: design and evaluation through simulation. *In* J.C. Thomas and M.L. Schneider (eds.), Human Factors in Computing Systems. Norwood, NJ: Ablex Press.

Shneiderman, B. 1997. Designing information-abundant websites: issues and recommendations. International Journal of Human-Computer Studies 47(1):5-30.

Silverstein, L.D. Human factors for color display systems: concepts, methods, and research. *Pp*. 63-82. *In* H.J. Durrett (ed.). Color and the Computer. Boston: Academic Press.

Smith, S.L., and N.C. Goodman. 1972. Another look at blinking displays. Human Factors 14(4): 345-347.

Smith, S.L., and N.C. Goodman. 1971. Blink coding for information display. Human Factors 13: 283-290.

Smith, P.A., I.A. Newman, and L.M. Parks. 1997. Virtual hierarchies and virtual networks: some lessons from hypermedia usability research applied to the World Wide Web. International Journal of Human-Computer Studies 47(1): 67-96.

Snowberry, K., S.R. Parkinson, and N. Sisson. 1985. Effects of help fields on navigating through hierarchical menu structures. International Journal of Man-Machine Studies 22: 479-491.

Snowberry, K., S.R. Parkinson, and N. Sisson. 1983. Computer display menus. Ergonomics 26(7): 699-712.

Snyder, H.L., J.J. Decker, C.J.C. Lloyd, and C. Dye. 1990. Effect of image polarity on VDT task performance. *Pp*. 1447-1451. *In* Proceedings of the Human Factors Society 34th Annual Meeting, Vol. 2. (Orlando, FL, October 8-12, 1990).

Spool, J., T. Scanlon, W. Schroeder, C. Snyder, and T. DeAngelo. 1997. Web Site Usability: a Designer's Guide. North Andover, MA: User Interface Engineering.

Staggers, N. 1993. Impact of screen density on clinical nurses' computer task performance and subjective screen satisfaction. International Journal of Man-Machine Studies 39:775-792.

Stanton, N.A., R.G. Taylor, and L.A. Tweedie. 1992. Maps as navigational aids in hypertext environments: an empirical evaluation. Journal of Educational Multimedia and Hypermedia 1(4): 431-444.

Tinker, M.A. 1955. Prolonged reading tasks in visual research. Journal of Applied Psychology 39: 444-446.

Tullis, T.S. 1988. Screen design. *Pp*. 377-411. *In* M. Helander (ed.). Handbook of Human-Computer Interaction. Amsterdam: North-Holland.

Tullis, T.S., J.L. Boynton, and H. Hersh. 1995. Readability of fonts in the Windows environment. Online version of poster presented at CHI '95 Conference. http://www.acm.org/sigchi/chi95/Electronic/documnts/intpost/tst_bdy.htm (February 9, 1998).

Vartabedian, A.G. 1971. The effects of letter size, case, and generation method on CRT display search time. Human Factors 13(4): 363-368.

Waters, C. 1997. Universal Web Design. Indianapolis, IN: New Riders Publishing.

Wixon, D., and J. Ramey (eds.). 1996. Field Methods Casebook for Software Design. New York: John Wiley & Sons, Inc.

Appendix. Some Web Design Guidelines

Many web design guidelines exist. Instone (1998) had indexed 27 different web sites containing web design guidelines as of March 31, 1998. A list of all 27 links is available at the Usable Web site (http://usableweb.com).

Some of the best-known guidelines are:

- Yale C/AIM WWW Style Manual. http://info.med.yale.edu/caim/manual/contents.html (March 31, 1998).
- Improving Web Site Usability and Appeal. http://www.microsoft.com/plan/IMPROVINGSITEUSA.HTM (March 31, 1998).
- Apple's Web Design Guide. http://applenet.apple.com/hi/web/web.html (March 31, 1998).
- IBM Web Design Guidelines. http://www.ibm.com/IBM/HCI/guidelines/web/web_guidelines.html (March 31, 1998).
- Creating Killer Web Sites. http://www.killersites.com/index.html (March 31, 1998).