The Effect of Heading Frequency on Comprehension of Online Information: A Study of Two Populations

LAURA D. SCHULTZ AND JAN H. SPYRIDAKIS

INTRODUCTION

For Web-savvy technical communicators, it is no surprise that online health information is an intrinsic part of the Web these days. A July 2003 survey (Pew Internet and American Life Survey) reported that 80% of adult Internet users—about 93 million American healthcare consumers (HCCs)—have searched for at least one of 16 major health topics online. Contributing to the growth of online health information is the current economics of today's healthcare system in which patients find themselves more responsible for their own healthcare.

As doctors spend less time in each patient appointment, HCCs often do not have their informational needs met. Reduction of access to healthcare providers (HCPs) is also apparent in the results of a 1998 MediMedia Managed Care survey that showed that 30% of family physicians were no longer accepting new patients. Further complicating the issue of diminishing access to HCPs, the legions of HCCs are growing as the baby boomer population ages for children and aging parents, and will soon be part of the graying population itself. The strain on access to medical care will continue to force people to be proactive in educating themselves about health and medical concerns (Deloitte Consulting and Deloitte and Touche 2001). As a result, more and more HCCs will turn to the Internet in search of healthcare information.

But more important than the growing reliance on and need for access to online health information is what consumers do as a result of accessing such information. According to Deloitte Consulting and Deloitte and Touche, of the HCCs who accessed medical information online, 30% said that they visited a doctor as a result, 43% talked to their doctor about a prescription, 47% encouraged a family member to see a doctor, and 42% made a treatment decision (1999). In a study by Fox and Ratnie (2000), 70% of the respondents reported using information obtained from the Web when making treatment decisions.

As both the amount of information online and the number of users accessing the information rapidly increases, designers of Web sites critically need empirical evidence to support the many design decisions they must make. One such design decision concerns how online prose should be structurally segmented—more specifically, how frequently headings should be placed to adequately signal readers about text structure.

Empirical research on headings in print reveals that headings facilitate comprehension in many ways. We know, for example, that "signs" in print text (such as headings and topic sentences) help readers perceive the organization of the text and thus create a schema that will help them comprehend the material. Intriguing questions remain, however, about how headings affect readers in online environments. For example, if headings help readers instantitate text schemata, emphasize the structure and topics of a text, and facilitate recall, could there be an optimum frequency of headings? And how would heading frequency relate to readers' perceptions of the text—that is, would readers perceive that they had learned more after reading information that contained headings? Such questions have not been examined in either print or online environments. But because the Web is becoming a primary information source for all types of readers, the effect of heading frequency in an online environment needs specific, empirical study.

The goal of the two studies reported here was to examine the effect of heading frequency on comprehension of online medical information with participant samples drawn from two different populations, one with a vested interest in obtaining medical information and one with a more neutral relationship to obtaining medical in-

Manuscript received 17 September 2003; revised 22 April 2004; accepted 26 April 2004.
formation. Although this study used online medical information as a test bed for an informational Web site, we believe that the study findings are generalizable to other types of informational Web sites.

LITERATURE REVIEW
Before presenting our study, we briefly review relevant literature concerning (1) the development of schemata through headings in print, (2) the effect of signals on readers, and (3) text and reader characteristics known to affect readers' use of signals.

Schemata: Effects on learning and retrieval of information
When learning new information, people create mental frameworks or schemata that help them to encode information and later retrieve it from long-term memory (Bartlett 1932; Anderson and Pearson 1984; Kintsch and Yarbrough 1982; Olthausen and Rolier 1988; Brooks and Damsereau 1983; Williams 1994). These schemata consist of mental outlines of known information that is often arranged hierarchically in a framework but is not necessarily taxonomic. New information can be attached or discarded if it does not seem to fit with the already accepted information—experiences and new information can change, adjust, correct, or fill in gaps in an individual's personal schemata (Williams 1994; Anderson and Pearson 1984). Thus, a schema is a grouping of related ideas or concepts.

Without a well-developed schema on which to add new information, readers cannot retrieve new information from memory. The steps involved in developing schemata can be viewed as a cycle of activation, comparison, approval or rejection, and instantiation.

1. When a schema is activated, it is mentally recalled to help interpret a situation or environment.

2. Comparison occurs when the new information from a situation is reviewed and checked against the existing, accepted information.

3. New information that is inconsistent or incompatible with the existing schema is rejected. However, if the new information is carefully reviewed and found to be credible and convincing, the existing schema may be adjusted (Anderson and Pearson 1984). One can think of such schemata as having various "in-boxes" or "nodes" for new information.

4. When the new information is accepted, nodes are instantiated with new information. New information can be added only when it relates to the schema.

It is important to note that the existing schema must be stable before new information can be added.

The research on schemata has important implications for writers who must use whatever tools they can to help readers activate appropriate schemata. Although one may presume that schemata are influenced by signals such as headings in an online text, this educated presumption needs to be empirically tested.

Text signals: Helping readers create schemata and recall information
Signals in text help readers build an accurate and useful schema for incoming text by emphasizing text structure and topics; the schema in turn facilitates recall (Lorch 1985; Meyer 1975). A text signal is information that is embedded in text but that does not actually add new information—it instead highlights a text’s structure, important points in a text, or both (Meyer 1984). Text signals can take many forms, including font changes, lead-in sentences, overviews, summaries, logical connectives, or headings (Lorch, Lorch, and Inman 1993; Lorch 1985; Meyer 1975; Spyridakis 1989a; Spyridakis 1989b; Spyridakis 1991).

Lorch and Lorch (1996b) found headings specifically to be important in helping readers create a schema.

Headings, in particular, explicitly communicate the topic structure of a text, thus providing support for readers’ attempts to construct a representation of the topic structure. (p. 262)

Headings can also signal specific information about text content and thus provide cues to the important content of the text (Eysenck and Keane 1990; Hartley 1987; Waller 1982).

Signals have been shown to affect how much and what type of information readers recall. In many studies, organizational signals led to recall of more text topics than when the organization was not signaled (Lorch and Lorch 1985; Lorch and Lorch 1995; Lorch and Lorch 1995a; Lorch and Lorch 1996a; Lorch, Lorch, and Inman 1993; Lorch 1985) found that signaled text (that is, text with preview sentences) significantly aided recall. Much research (Spyridakis 1989a; Spyridakis 1989b; Spyridakis 1991; Lorch, Lorch, and Matthews 1985; Lorch and Lorch 1986; Lorch and Lorch 1995; Lorch and Lorch 1996a; Lorch, Lorch, and Inman 1993; Hartley, Kenely, Owen, and Trueman 1980; Sanchez, Lorch, and Lorch 1993) implies that signals are helpful for text recall because they (1) draw the reader's attention to important areas of text, (2) facilitate schema construction for that text, and (3) direct the order of text recall. Further, Spyridakis (1989a, 1991) points out that some signals have stronger effects with inferential versus factual comprehension and with delayed versus immediate testing.

It is important to note that signals (such as headings) can affect the organization of recall (Lorch 1993; Lorch, Lorch, and Matthews 1985). Research has shown that when organizational signals are present, the recalled content seems to be organized as it was in the text (Lorch and
Lorch 1996b). Many studies have supported this “levels effect,” which proposes that information at a higher hierarchical level in the text will be better recalled than information at lower levels (Kintsch and colleagues 1975; Mandler and Johnson 1977; Meyer 1975; Meyer 1977; Meyer 1985; Britton and colleagues 1980; Kintsch and Yarborough 1982).

Although considerable research in print-based text has shown that headings help subjects to retain and retrieve information, little is known about the effect of headings in online text. Lorch and Lorch (1986) studied the effect of summary sentences that approximated sentences in an online environment and found that subjects remembered text marked by such headings better than text without. However, this environment consisted of a video display screen with one sentence replaced by the next each time the space bar was pressed, an environment suitable for that experiment but one that does not resemble today’s online Web environment.

On the usability side, Jakob Neilsen (1997) encouraged the use of headings in online text to help readers scan text—but once again, this advice is not based on empirical research with naturally occurring Web pages. Research is needed to ascertain the effect of headings on comprehension of online text, as well as the optimal frequency of headings to facilitate comprehension in online environments.

Text and reader characteristics: Mediating readers’ use of signals
In addition to the effects of signals in general, some specific text and reader characteristics can affect readers’ use of signals. In particular, readers’ perceptions of text difficulty or familiarity can affect how much they rely on signals. Further, readers’ interest in and enjoyment of text can affect their use of signals.

To elaborate, if text is relatively difficult to understand, signals, including headings, may be more likely to help readers comprehend the material (Spyridakis 1991). In contrast, easy text (that is, text without a lot of hierarchical levels or topics) shows less benefit from the use of signals (Lorch and Lorch 1996a; McDaniel, Einstein, and Lollis 1988).

Also, headings seem to help more with recall of unfamiliar topics than with familiar topics. Research suggests that familiar topics are already encoded with relevant schema and that readers may not need text signals to assist in recall (Blasko and Connine 1993; Conway and Dewhurst 1995; Sadoski, Goetz, and Fritz 1993; Niederhauser and colleagues 2000; Spyridakis 1991). But unfamiliar topics may require extra assistance:

Even if the unfamiliar topic is identified and recognized as important to the text’s structure, it may be more difficult to organize what is said about the topic and relate it to a common context in memory. In this situation, introducing the topic with a heading may provide several benefits that translate into greater accessibility at subsequent recall. (Lorch and Lorch 1996b, p. 263)

It is possible that experienced readers (such as adults) can encode information regardless of the presence of signals in print text (Lorch and Lorch 1985; Lorch and Lorch 1995; Lorch, Lorch, and Inman 1993). One hypothesis suggests that practiced readers are already aware of the general schema of expository text and know that it should have a topic structure. Thus, they can create their own mental outline as they read. Nonetheless, signals can still help these experienced readers in developing accurate and complete representations of the text.

The use of headings has also been shown to increase readers’ interest in text (Lane, Newman, and Bull 1988), which in turn has been shown to lead to higher comprehension (Asher 1980). Tying interest and familiarity together in a causal model, Sadoski, Goetz, and Fritz (1993) found that familiarity helped increase comprehension, and comprehension in turn created a state of “interestengness,” which increased immediate recall of information.

Clearly, interest, familiarity, text difficulty, and comprehension are interrelated. And if readers are more familiar with or interested in a text, or if they find it relatively easy, factors that by themselves can facilitate comprehension, then headings might prove to be of less benefit.

Research also suggests that the more enjoyable the reading activity, the better the reader’s performance. For example, the more a reader enjoys reading a text, the higher the comprehension level (Marchese 1998). Thus, headings may be less vital for readers if reading a text is enjoyable and they are easily able to absorb information. Of course these reader and text variables start to inform each other in that a more interesting, familiar, or enjoyable text may be perceived as easier to read—all scenarios that might lead to less reliance on text signals in order for readers to form appropriate text schemata. The interrelationship of these text and reader variables with signals such as headings is investigated in the study presented in this article.

Hypotheses and research design
The study presented here investigates the effect of heading frequency on comprehension. (This study also examined the effect of unmediated online health information on participants’ emotional states, but those results are reported in Schultz and Spyridakis [unpublished].)

The review of the literature led to three hypotheses.

H1. Headed text will result in better comprehension than nonheaded text.
H2. The more frequent the headings, the higher the scores on a comprehension test.

H3. Readers’ perceptions of text familiarity, difficulty, interest, and enjoyment will relate to the effect of headings on comprehension.

These hypotheses were investigated through two experiments that assessed the effect of heading frequency on comprehension of two experimental texts with two different samples:

- Adults from the community at large, many of whom have medical needs related to the experimental texts
- University students, most of whom would have no immediate need for the information in the experimental texts

We presumed that these two samples would have differing perceptions of the texts (for example, the adults would be more familiar with and perhaps more interested in the text than the students).

The studies included two between-subjects variables:

- Frequency of headings (no headings, approximately 100 words per heading, approximately 200 words per heading, approximately 300 words per heading)
- Web page topic (osteoarthritis or rheumatoid arthritis)

Dependent measures reported in this article included comprehension, and perceptions of the Web pages and knowledge learned.

EXPERIMENT ONE

Methods

Participants One hundred and fifty-one individuals volunteered to participate in the first study, ranging in age from 18 to 82 years (mean = 45.54, standard deviation = 13.80). Participants were recruited through a variety of venues that targeted adults who in the course of their daily lives might seek out medical information. One venue was an arthritis Web site, the Arthritis source. Other recruiting venues included flyers placed in various doctors’ offices, medical clinics, and a senior center.

Materials—arthritis texts The materials for the study consisted of two Web sites on arthritis, two surveys, and instructional pages. Content for two diseases (osteoarthritis OA and rheumatoid arthritis RA) was used to ensure the stability of the results and eliminate the chance that results would be content specific. The topics of OA and RA were chosen with a goal that participants would find these two experimental texts to be moderately difficult, interesting, familiar, and enjoyable to read.

The two texts were adapted from Web pages about OA and RA on the Arthritis source Web site. The Arthritis source (http://www.orthop.washington.edu/arthitis) is a comprehensive medical Web site created by the Washington State chapter of the Arthritis Foundation and two academic units at the University of Washington (the Department of Orthopaedics and Sports Medicine and the Bone and Joint Center). The site is authored by physicians, formatted by document designers, and tested for usability by the Arthritis source team. These texts gave a basic introduction to the diseases, treatments, and prognoses.

Four versions of each arthritis text were created—one for each heading frequency level, resulting in eight Web pages. To avoid confounding the heading frequency variable, only one hierarchical level of headings was used in the experiment (that is, no subheadings were used). The levels of heading frequencies were determined by reviewing many medical Web sites and choosing commonly found heading-to-word-count ratios:

- No headings (a control condition)
- High frequency of headings (one heading approximately every 100 words)
- Medium frequency of headings (one heading approximately every 200 words)
- Low frequency of headings (one heading approximately every 300 words)

Headings were created by constructing new headings or editing existing headings. In general, these headings consisted of noun or participial phrases that were between one and seven words long (mean = 3.02, standard deviation = 1.38). The goal in designing the headings was to summarize the main point of the paragraph or paragraphs that followed each heading.

Four technical writers rated the headings for how well they represented the associated information. If the raters believed the headings were inaccurate, they rewrote the headings, which were then revised. Across the four heading conditions, the eight Web pages had approximately the same number of words (OA pages average word count = 1,841, standard deviation = 55.0; RA pages average word count = 2,120, standard deviation = 7.05).

The Web pages for the OA and RA texts were formatted so as to reproduce the Arthritis source Web site’s format. The only changes to the look and feel of the Web pages were the modification of each text into one continuous Web page per topic, the removal of the table of contents from the left navigation bar, the removal of embedded links, and the addition of brief instructions and a single navigation button that took participants to Survey II (see Figure 1). No graphics or illustrations accompanied the text.

Materials—instructional pages There were several instructional screens in this experiment. The welcome page, which overviewed the study, told participants that the experiment would take less than half an hour, that they
Figure 1. Example of the layout of arthritis Web text.

would remain anonymous, and that by clicking “Continue to Experiment,” they were consenting to participate in the study and verifying they were over 18.

The next Web page asked participants whether they or a friend had OA or RA, or whether they were just curious about the study, a question that was important for the emotional status portion of the experiment (Schultz and Spyridakis unpublished) The next instructional page provided a “Continue to Survey II” link, yet gave participants the opportunity to click the Back button if they were not finished reading the arthritis page. A final instructional page at the end of the experiment provided an opportunity to enter a drawing for an Amazon.com gift certificate.

Materials—surveys The experiment used two surveys: one presented before the OA and RA arthritis texts, and one after the texts. Survey I contained demographic questions about participants’ use of medical information sites, the Web, and the Internet in general; their motivation for participating in the study; and their education. The demographic questions were informed by the Georgia Institute of Technology’s 10th WWW User Survey (Graphic, Visualization, and Usability Center 1999). Survey I also presented a scenario telling participants that if they were not researching arthritis for personal reasons, they should imagine that they were trying to find information about arthritis to help a relative or friend learn more about the effects of the disease and best treatments. Survey I then asked participants to rate their emotional status concerning their health regarding arthritis or their friend’s or relative’s health regarding arthritis.

Survey II asked participants their age and gender and inquired whether they were native English speakers. It also asked participants to rate how interesting, difficult, and enjoyable the information in the arthritis text was; how familiar they were with it; and how much new knowledge they felt they gained. Survey II reiterated the emotional status questions, and asked about participants’ comfort level that their or their friend’s arthritis could be managed.

Survey II also contained 15 multiple-choice compre-
hension questions about the arthritis texts (different question sets for OA and RA texts), each with three possible answers. Comprehension questions were developed following the guidelines in Writing test items to evaluate higher order thinking (Haladyna 1997). The comprehension questions were approved by Dr. Fredrick Matsen (an Arthritis source author) and then tested for passage dependency with 61 technical communication students who did not read the arthritis texts—the goal being to obtain an equal distribution of answers across all possible answers per question. The questions were edited and retested until all the questions met the passage dependency goals.

Pilot tests of the experimental Web site revealed that pilot subjects could navigate through the site and complete the experiment without any difficulty or confusion.

Overall procedure Participants accessed the experiment by following online links on the Arthritis source or entering the URL found in flyers. They read the general instructions for the experiment and selected the “Continue to Experiment” link. Participants then identified their connection, if any, to arthritis, took Survey I, and were randomly assigned to an arthritis text (participants who identified a relationship with one of the two types of arthritis received one of the texts on the relevant arthritis type). After an arthritis text Web page opened, participants were instructed to read the arthritis content at their leisure, and when finished, click the button that would direct them to Survey II. Once participants submitted their answers to Survey II, the gift certificate screen opened and invited them to enter the drawing. All data was analyzed using SPSS 11.5 software.

Results and discussion of Experiment One
After a review of demographic results, we discuss the comprehension results. Only results at an alpha level ≤ .05 are discussed as significant. An alpha level of .05 would mean that there is only a 5% chance that the findings reported as significant result from chance (when p values—probability values—that are reported with individual statistics are equal to or less than the specified alpha level, the finding conveyed by the statistic is deemed significant).

Of the 151 participants in this experiment, 76 participants read the OA texts and 75 read the RA texts. Given our expectation that the two texts would not yield different results, the texts are collapsed for analysis of all variables unless significant text differences were found, in which case appropriate statistics were run to reveal the interactions. Given the subject pool for Experiment One, we deemed it wise to check for text differences.

Demographic results As expected, given the way in which participants were recruited, most participants (95%) were native English speakers. Additionally, participants were fairly well educated, with almost 62% (n = 84) having at least a bachelor’s degree and 17% (n = 26) identifying themselves as medical professionals (such as nurses, operating room technicians, pharmacists). Of those participants who identified their gender, 23 were males and 123 were female.

As one might expect of a self-selecting population that participates in an online study, the participants seemed to be fairly experienced Internet users. The average time spent on the Internet per week was 11.89 hours (standard deviation = 10.98). Given participants’ relatively high hourly usage of the Internet weekly, it was not surprising that participants rated themselves as extremely comfortable with computers (mean = 4.15, standard deviation = 1.232; 1 = very uncomfortable; 5 = very comfortable) and with the Internet (mean = 4.19, standard deviation = 1.240). These variables were highly correlated (r[149] = .911, p = .000). Over 75% of participants reported that they used the Web to search for general information between a few times a week and daily (see Table 1). In contrast, only 35% searched for health information a few times a week or daily.

Enjoyment, interest, familiarity, and difficulty Our goal in selecting the experimental texts was to use texts in which participants might be somewhat likely to employ headings to help them learn and recall new information. If participants found the texts to be extremely enjoyable,
interesting, familiar, or easy, then headings might be of little benefit. Thus, analyses were run concerning participants’ ratings of their (1) enjoyment of, (2) interest in, (3) familiarity with, and (4) perceived difficulty of the texts. Although no differences were anticipated between texts, t tests were run to check for text differences. The tests revealed no differences between texts for enjoyability and interest, so these variables were further analyzed across texts. Participants rated the texts close to the midpoints on the five-point enjoyability scale (mean = 3.17, standard deviation = .895; 1 = very unenjoyable, 5 = very enjoyable), and slightly above the midpoint on the five-point interest scale (mean = 3.53, standard deviation = 1.575; 1 = very uninteresting, 5 = very interesting).

Because t tests to assess text differences on familiarity and difficulty revealed significant differences, 2 (text) x 4 (heading frequency) analyses of variance (ANOVAs) on familiarity and difficulty were conducted to assess the effect of these variables on the dependent variable of interest with one statistical routine. There was a significant main effect for text on familiarity ($F(1,142) = 13.823, p = .000$, partial $\eta^2 = .089, \alpha = .958$) with readers of the RA text being more familiar with the information (mean = 3.72, standard deviation = 1.07; with 1 = very unfamiliar, 5 = very familiar on a five-point scale) than readers of the OA texts (mean = 3.04, standard deviation = 1.14).

The effect size ($\eta^2$), which reveals the strength or meaningfulness of the effect, reveals that the effects were substantial, given that Cohen (1988) defines a medium effect as at least .06 and a large effect as at least .14, and the text on familiarity resulted in a respectable effect size of .089. This result also had an incredibly high power (1) of .958, with .80 being a common cut-off point. Power is an important concept in that it reveals the probability of finding an effect, if one actually exists—in this case, the power of .958 means that there was a 95.8% chance of finding an effect in the true presence of an effect. In general, results with low power may be caused by small sample sizes or very small effect sizes.

The two-way ANOVA to assess the effect of text and heading condition on difficulty revealed a significant difference between the two texts ($F(1,141) = 5.097, p = .026$, partial $\eta^2 = .035, \alpha = .611$), with readers of the RA text perceiving the information to be easier (mean = 3.23, standard deviation = 1.28; with 1 = very difficult, 5 = very easy on a five-point scale) than readers of the OA text (mean = 2.82, standard deviation = .10). The effect was rather small (Cohen 1988 defines a small effect as .10) and had a low power of only .611—a result that would suggest that if the sample had been larger, the effect might have been even larger.

The higher ratings of familiarity and text easiness for readers of the RA text might have reduced the possible benefits that headings could provide because these participants would be more able to depend on their prior knowledge and reading ability and thus would not be in as much need of help from textual signals. The difference in participants’ familiarity with the RA and OA texts may relate to the ramifications of having a personal connection to RA versus OA. Because RA may be more disruptive to one’s lifestyle, people with a connection to RA may be more proactive about learning about the disease, and thus read more about it, and find it easier to read and more familiar.

Of specific interest to the study goals, we also found a significant difference between heading frequency conditions on difficulty ($F(3,141) = 7.722, p = .000$, partial $\eta^2 = .141, \alpha = .987$). The effect size ($\eta^2$) for text was quite large, and this finding had substantial power of .987, suggesting that there was a meaningful effect for heading frequency on perceived difficulty of the two texts. A follow-up test (Scheffé post hoc test), used to determine which conditions differed, revealed that the no-heading condition was deemed significantly easier than all other heading conditions ($p < .05$) (see Table 2).

The fact that the readers who saw the no-headings condition judged the texts to be significantly easier than readers who saw any of the headlined texts was unexpected. An examination of the texts as they appeared on the Internet (Figure 1) revealed that many paragraphs were quite short, often with only one or two lines per paragraph and with a line of space between paragraphs (on a 17” monitor with Internet Explorer set on medium text size). Because the texts were not “copy dense” to begin with, readers may have perceived the no-headings texts to be quite sufficient in terms of the amount of structure revealed.

### TABLE 2: DIFFICULTY RATINGS BY HEADING FREQUENCY

<table>
<thead>
<tr>
<th>Heading Frequency</th>
<th>Mean (1 = very difficult, 5 = very easy)</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High (n = 42)</td>
<td>2.69</td>
<td>1.12</td>
</tr>
<tr>
<td>Medium (n = 37)</td>
<td>3.00</td>
<td>1.25</td>
</tr>
<tr>
<td>Low (n = 34)</td>
<td>2.71</td>
<td>1.06</td>
</tr>
<tr>
<td>None (n = 36)</td>
<td>3.75</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Comprehension results. To assess the effect of the heading frequency on comprehension, we first ran a t test to check for text differences between texts—none were found; participants correctly answered about two-thirds of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The comprehension results did not confirm the hypothesis that frequent headings would lead to higher comprehension scores. The lack of significant differences between any of the heading conditions may relate to participants’ ratings of the text. Participants’ interest in the texts significantly correlated with their total comprehension ($r(149) = .158, p = .053$) and their familiarity inversely correlated with comprehension ($r(148) = -.319, p = .000$). Further, participants’ interest ratings also positively correlated with perceived new knowledge gained ($r(149) = .153, p = .062$). In other words, readers who were more interested and less familiar with the text comprehended more, and participants who were more interested believed they had gained more new knowledge.

**EXPERIMENT TWO**

**Methods**

**Participants** For this experiment, 112 participants volunteered to participate in this study, ranging in age from 18 to 32 years ($mean = 21.95$, standard deviation = 3.71). Participants were recruited through two technical communication courses required for students in the College of Engineering at the University of Washington.

**Materials** The materials for the study consisted of the same materials used in Experiment One with an additional Web page that allowed participants to register for participation credit in their classes. Students were also asked their year in school.

**Overall procedure** The overall procedure was the same as in Experiment One. Participants received class participation points and could also enter a drawing for an
TABLE 4: WEB USE FREQUENCIES FOR STUDENT SAMPLE

<table>
<thead>
<tr>
<th>Use of Web to Search for:</th>
<th>Never</th>
<th>Less Than Once A Month</th>
<th>A Few Times A Month</th>
<th>A Few Times A Week</th>
<th>Daily</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Information (n = 112)</td>
<td>0.9%</td>
<td>0.9%</td>
<td>8.0%</td>
<td>29.5%</td>
<td>60.7%</td>
</tr>
<tr>
<td>Health/Medical Information (n = 112)</td>
<td>15.2%</td>
<td>58.9%</td>
<td>22.3%</td>
<td>2.7%</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

Amazon.com gift certificate. All data was analyzed using SPSS 11.5 software.

Results and discussion of Experiment Two

After a review of the results of the demographics questionnaire, we discuss the comprehension results. Only results at an alpha level ≤ .05 are discussed as significant. Fifty-four participants read the OA texts, and 58 read the RA texts.

Demographic results

Of the participants who responded to the question about native language, 74.1% (n = 83) identified themselves as native English speakers. No participants identified themselves as medical professionals. Of the participants who identified their gender, 86 were male and 25 were female. These students were almost evenly distributed in thirds across their sophomore (29.5%), junior (26.8%), and senior years (29.5%) in college.

The participants seemed to be experienced Internet users, spending on average 16.89 hours weekly (standard deviation = 14.49) and significantly more time per week on the Internet than the participants in the Experiment One (t (258), 3.144, p = .002). As expected, participants rated themselves as very comfortable with computers (mean = 4.08, standard deviation = 1.467; where 1 = very uncomfortable, 5 = very comfortable) and with the Internet (mean = 4.06, standard deviation = 1.490).

Over 90% of participants reported that they used the Web to search for general information between a few times a week and daily (see Table 4). As might be expected, only 3.6% of the student participants reported accessing medical information on the Internet a few times a week or daily.

Enjoyment, interest, familiarity, and difficulty

The next analyses concerned participants’ ratings of their (1) enjoyment of, (2) interest in, (3) familiarity with, and (4) perceived difficulty of the text read. Because t tests revealed no differences between texts on these variables, further analyses were conducted across texts. Participants rated the text slightly below the midpoint on the five-point enjoyability scale (mean = 2.85, standard deviation = .796; where 1 = very unenjoyable, 5 = very enjoyable) and near the midpoint on the five-point interest scale (mean = 3.19, standard deviation = 1.053; where 1 = very uninteresting, 5 = very interesting). The students’ ratings on both the enjoyability and interest scales were significantly lower than the adult participants’ ratings in Experiment One: enjoyability t (261) = 3.054, p = .003; interest t (261) = 2.199, p = .029.

The students rated their familiarity quite low on the familiarity scale (mean = 2.29, standard deviation = .907; where 1 = very unfamiliar, 5 = very familiar) yet they perceived the text as neutrally difficult (mean = 3.17, standard deviation = .826; where 1 = very difficult, 5 = very easy). The student participants were significantly less familiar with the material than the adult participants (t (260) = 8.221, p = .000), but this lack of familiarity is to be expected since the students were on average 22 years old and would not be as likely to be as familiar with the pathology of OA or RA as the adult participants. A t test for the effect of study population on text difficulty was not significant.

Comprehension results

Because the student sample had such a high percentage of nonnative English speakers, the comprehension results were first analyzed to determine whether native versus nonnative speakers would differ. The t test on the effect of native language on comprehension was significant (t (110) = 2.909, p = .004), with nonnative speakers answering significantly fewer questions correctly (mean = 7.10, standard deviation = 3.44, n = 29) than native speakers (mean = 8.88, standard deviation = 2.60, n = 83). Therefore, nonnative speakers were removed from analyses of comprehension and perceptions of new knowledge gained.

As expected, a t test to assess the effect of text on comprehension revealed no significant difference between texts (mean for OA = 8.47, standard deviation = 2.86;
TABLE 5: COMPREHENSION SCORES BY HEADING FREQUENCY

<table>
<thead>
<tr>
<th>Heading Frequency</th>
<th>Comprehension Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High (n = 25)</td>
<td>8.12</td>
<td>2.56</td>
</tr>
<tr>
<td>Medium (n = 15)</td>
<td>10.33</td>
<td>2.50</td>
</tr>
<tr>
<td>Low (n = 20)</td>
<td>9.35</td>
<td>2.35</td>
</tr>
<tr>
<td>None (n = 83)</td>
<td>8.35</td>
<td>2.57</td>
</tr>
</tbody>
</table>

mean for RA = 9.19, standard deviation = 2.35); however, the students’ comprehension scores were significantly lower than the adults’ scores in Experiment One (t (232), 3.009, p = .003). A one-way ANOVA, run to assess the effect of heading frequency on comprehension, was significant: F (3, 79) = 3.043, p = .034, partial η² = .104, α = .05. The effect size (η²) of .104 was quite respectable, in that Cohen (1988) defines a medium effect as at least .06 (a large effect is defined as at least .14).

Further analysis (a post hoc LSD test) to assess which conditions differed from each other revealed that participants in the medium frequency heading condition had significantly higher comprehension (p = .009) than participants in the high frequency heading condition (see Table 5).

The issue of comprehension was further investigated by looking at participants’ ratings of perceived new knowledge gained. Because a t test revealed no difference between texts for perceived new knowledge gained (mean for OA = 3.14, standard deviation = .692; mean for RA = 3.34, standard deviation = .111), the effect of perceived new knowledge gained was analyzed across texts with a one-way ANOVA, which also was not significant. As with the adult participants, the more new knowledge that participants reported gaining, the more interested they were (r (81) = .303, p = .006).

As becomes apparent in the conclusion, there are some interesting similarities between the adult and student participants that are very telling about headings in online text and have implications for quantitative and qualitative research.

CONCLUSIONS AND FUTURE RESEARCH

There are several points that communicators should note as research into online communication becomes more sophisticated. This Internet-based study delivered to users in naturalistic settings sets the stage for further research of this type. And the use of a health information topic makes examining two types of readers—one with a vested interest in the information and one with less of a connection—extremely fruitful.

To begin with, the comprehension results from the two experiments reveal an intriguing difference between the two samples. Although we had hypothesized that more frequent headings in an online environment would result in higher comprehension scores, only the student sample showed differences on comprehension among heading frequency conditions, with students in the medium-frequency heading conditions comprehending more than students in the high-frequency conditions. Interestingly, though, the adult sample believed they had gained significantly more knowledge from the medium-frequency heading condition than from the high-frequency condition.

It is noteworthy that actual and perceived comprehension tended to favor the medium-frequency heading condition (a heading approximately every 200 words). This medium frequency of headings apparently helped the student participants instantiate a text schema that in turn facilitated their comprehension, and the adult participants may have had a related experience that influenced their perceptions of new knowledge gained. In both cases, the high-frequency headings (a heading approximately every 100 words) appeared to impede either comprehension or perceptions of new knowledge gained.

With both the adults and the students, the comprehension and perceived knowledge results are informed by the participants’ backgrounds and text perceptions. First, all participants were experienced Internet and Web users, and although the adults were more educated and visited health information Web sites more often than the students, the students spent more time on the Internet in general.

These online readers may have been very familiar with the common convention of relatively short paragraphs in online text, a principle that Spool (1999) strongly advocates. We cannot help but wonder whether the participants’ high level of Internet experience may have interacted with the brevity of the paragraphs in the experimental texts. Technical communicators can take note that other textual situations may require headings to make the text comprehensible. In online texts that do not follow the convention of short paragraphs, or that have difficult or complex information, headings may be very helpful to readers.

Further, adults had more personal experience with either RA or OA than the students. This experience with arthritis information was further supported by the adults’ view of text difficulty and familiarity versus the students’ views, and with the adults’ finding the RA text to be more familiar and easier than the OA texts. This finding was also reinforced by adult RA readers who reported more new knowledge gained than adult OA readers.

The fact that no text differences occurred for the stu-
The Effect of Heading Frequency on Comprehension of Online Information

Schultz and Spyridakis

students is not surprising in that the students across the board were less familiar with the texts and rated them to be less interesting and less enjoyable than the adults. This situation positioned the students to gain the most from headings—and it was, in fact, the students whose comprehension benefited from one of the heading frequency conditions. Taking this finding into consideration, designers of Web information may find it wise to include a moderate number of headings in text that may be unfamiliar to readers. Further, when information is especially important for readers to comprehend but might be a bit dry and technical, a moderate number of headings may help to increase learning even when the readers are not particularly interested in the topic.

The difference between the actual comprehension and perceived comprehension results speaks to a methodological consideration that researchers and usability experts should consider—the need to gather both quantitative performance data as well qualitative perceptual data. Although it is vital to consider reader performance, especially when dealing with specific topics and subjects, it is also important to consider reader perceptions of performance.

Readers who perceive that they have learned new information from a Web site will be likely to return to the Web site, a situation that can further increase actual comprehension and give a Web site every opportunity to educate. It may be wise for communicators to query readers about the amount of new information they feel they have learned after reviewing a Web site. Such questions may give online information designers valuable insights into actual versus perceived learning.

A further methodological consideration concerns the type of comprehension assessed. We have wondered whether the comprehension test in the studies presented here was too focused on factual comprehension as opposed to inferential comprehension, the latter of which has been shown in the literature to be more positively influenced by headings. This inadvertent focus on factual questions occurred because of the multiple revisions during the process of ensuring passage dependency and obtaining expert approval of the comprehension questions.

The studies presented here begin to illuminate both the need for more empirical study and the wealth of information that may be uncovered by researching the effect of text design variables in online environments. Future research is needed to flesh out the nuances of the effects of heading frequency in an online environment. Future studies might include other text topics, perhaps a topic that is somewhat unfamiliar or less interesting to readers and that might cause them to depend more on headings to build appropriate schemata.

Future research might also use longer texts with longer paragraphs, which may again require readers to depend more on headings. The hierarchy of headings might also be varied in future research—perhaps multiple layers of headings would communicate a clearer structure and hierarchy to participants. Also, testing participants’ comprehension with inferential and factual tests, and retesting after an interval (for example, one to two weeks) might reveal further effects for heading frequency.

Comprehensible information is an indispensable need for information seekers who are filling in their knowledge gaps through their browser windows. Researchers have a great and intriguing opportunity to be on the front lines of a brand new area of empirical research, testing established design principles from the print environment in the online world. Further research may determine which principles work in both print and Web environments, and which are unique to either arena.

ACKNOWLEDGMENTS

The authors would like to thank the Arthritis source team for their gracious assistance and cooperation. We would especially like to thank Dr. Fredrick Matsen at the University of Washington Medical Center for his generosity in donating his time and expertise in the inception and creation of the materials for this research.

REFERENCES


MediMedia Managed Care. 1998. Survey of access to physician services. Yardley, PA: MediMedia Managed Care.
The Effect of Heading Frequency on Comprehension of Online Information


Pew Internet and American Life Survey. 2003. Internet health resources: Health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access. http://www.pewinternet.org/reports/toc.asp?Report=95.


LAURA SCHULTZ works at The Boeing Company as a technical editor, and is a senior member of the Puget Sound chapter of STC. She received her MS in technical communication from the University of Washington. Her areas of primary research are signaling and the communication of medical information online. Contact information: laura.d.schultz@boeing.com.

JAN H. SPYRIDAKIS is a professor of technical communication at the University of Washington. Her research focuses on document and screen design variables that affect comprehension and usability. Web design, cross-cultural audiences, and the refinement of research methods. She has won numerous teaching and publication awards, and she often teaches seminars in industry. Contact information: jansp@u.washington.edu.