The World Wide Wait: Effects of Delays on User Performance

Paula R. Selvidge
Software Usability Research Lab
Wichita State University

Barbara Chaparro
Software Usability Research Lab
Wichita State University

Gregory T. Bender
International Business
Machines Corporation

This study investigated the effect of web page download delays on user performance. Participants completed tasks on a major airline web site with varying download delays of 1 second, 30 seconds, and 60 seconds. The dependent measures included lostness, frustration, task success, and task difficulty. Delays did not affect lostness, but frustration was increased by the longer download delays of 30 and 60 seconds.

INTRODUCTION

Limited bandwidth is a persistent problem with the Internet. The 9th World Wide Web User Survey reported that speed, or taking too long to download pages, continues to be the number one problem reported by users (GVU web site, 1998). Users are also more likely to lose interest in a site if the download times exceed 10 seconds (Nielson, 1996). This problem translates into lost profits for e-commerce web sites. Unfortunately, addressing the factors that affect download times is not as simple as making improvements at the users' end, such as increasing connection bandwidth or speed of the users' modem and computer. The problems related to delayed response times are determined by many factors, such as bottlenecks due to the passage of large amounts of information, the number of users, the server's connection to the Internet, and hardware associated with the server (Nielson, 1997). All of these problems have a cumulative effect on the speed of response time on the Internet and addressing these issues requires considerable time and money. Implementing technological improvements to deliver greater bandwidth are not expected until the year 2003 (Nielson, 1997). In the interim, web sites should be designed to minimize download time.

Recent studies have investigated how perceptions of web sites are influenced by delays in retrieval time (Ramsay et al., 1998; Sears et al., 1997). Ramsay and colleagues (1998) examined the effect of delay length, two seconds to two minutes, on the perceptions of web pages while browsing. Seven different page types (scientific, business, advertisement, personal, history, instructional, entertainment) were presented at varying delay lengths and users rated the pages on how interesting the page content was and the difficulty with which the page could be scanned. The pages were also classified by style; text-only, graphics, with links, with few links, and balanced text and graphics. Significant differences were found between the pages with the shortest and longest delays, with differences starting at 41 seconds and longer (Ramsay et al., 1998). The pages with longer delays were rated as less interesting and more difficult to scan. No differences were found in ratings between different page styles, that is, all pages regardless of style received equally poor ratings in the longer delay conditions.

The effect of delays on perceived usefulness, organization, and quality of information on a web site has also been examined (Sears et al., 1997). Two versions of a web site (text-only and text plus graphics) were presented with a short, medium, or long delay length. The web pages were delayed different amounts of time to simulate real network delays, unlike the Ramsay (1998) study in which the delay lengths were fixed. The mean delays for each condition were .58 seconds (short), 3.5 seconds (medium), and 6.8 seconds (long). Users rated the ease of locating information, information organization, quality of information, and navigation problems. Users were sensitive to delays in both text-only and text plus graphics conditions. In contrast to the Ramsay (1998) study, page style did affect perceptions of the web pages. For text-only pages, shorter delays resulted in less favorable responses, and for text plus graphics pages, shorter delays were associated with more favorable responses.

Previous research has examined the effect of delays on preference and subjective evaluations of web sites, so the purpose of the current study was to examine the effect of web page download delays on user performance. More specifically, this study addressed the hypothesis that longer delays would increase lostness and frustration, and decrease task success. The hypotheses were: 1) 30 and 60 second delay conditions would increase lostness on the web site and decrease task success more than 1 second delays, and 2) frustration would be higher on 30 and 60 second delays than 1 second delay conditions.

METHOD

Participants

Fourteen undergraduate students, eight women and six men, from a large midwestern university participated in this experiment. The mean age for participants was 23.14 (SD = 6.24). Two requirements for participation in the study were: 1) participants must not have used the selected web site previously, and 2) participants must use the Internet at least five hours a month. The mean Internet usage was 9.6 hours (SD = 7.48). The students received extra credit points in their introductory psychology or upper-division psychology classes for participation in the experiment.

Materials

A Pentium-class computer with a 17" VGA monitor was used in the experiment. An offline reader program was used to download the airline web site, so the web site was accessed from the hard drive during the experiment in order to control download delays for page loading time. This web site was chosen based on prior usability studies, which showed that the site ranked high on preference and performance measures (Selvidge, 1999). A Microsoft Visual Basic 6.0 program was used to present web pages at specific fixed download delays and record navigation. The program utilized the Microsoft Internet Explorer 4.0 browser, and the interface provided Back, Forward, Stop, and Refresh buttons.

A pretest was conducted to ensure difficulty level of tasks and identify download delay times. The pretest examined the effect of delay (1, 10, and 20 seconds), on lostness, task success, frustration, and task difficulty. Results indicated no effect of delay for lostness, task success, frustration, or task difficulty. These findings suggested that perhaps the delays were not long enough, so the delay times were increased to 1, 30, and 60 seconds for the current experiment.

Design

A Randomized Block Design (Kirk, 1995), with delay length (1 s, 30 s, 60 s) varied within-subjects, was used to test the effect of download delays on user performance. Nine information retrieval tasks at the same difficulty level were presented, with three tasks at each delay condition. Participants were randomly assigned to delay conditions and the order of the tasks and delay conditions were counterbalanced.

Procedure

All participants were tested on an individual basis in a one-hour session. Participants completed an informed consent and demographic questionnaire assessing experience on the Internet. The participants were told that they could receive a cash prize of $20 if they had the fastest average time across all tasks. The instructions also stressed that they could quit on a task if they could not find the information, and it would not affect their chances of winning the cash prize. The tasks began at the airline web site home page. Each page on the web site loaded at the same fixed download delay throughout the four tasks. For example, if the participant was assigned to the 10 second delay condition, then each page loaded at 10 seconds. The delay began with the users' action of selecting a hyperlink, Back or Forward command button. During the delay period, the page where the user initiated the action remained on the screen while a message window (2.5 mm x .75 mm) with the text "Working" was presented in the center of the screen to provide feedback that the page was loading. After the delay, the hyperlink or page selected was presented immediately with all text and graphics fully loaded or visible. The user could continue to scroll the page, after selecting a hyperlink, until the delay was over and the new page was presented. The participants completed nine information retrieval tasks on the airline web site, with each delay condition presented three times. The experimenter cleared the browser's cache to reset visited links, so performance was not affected on subsequent tasks.

Performance measures. The performance measures were: 1) lostness measure and 2) success of locating task information. The lostness measure included a comparison between the number of nodes required to complete a task (optimal path) and the number of different nodes visited while searching for information (Smith, 1996). The optimal path or number of nodes required to complete the task was determined by the experimenter prior to testing. To calculate the lostness measure, the optimal path was divided by the number of different nodes visited. If the value was closer to 1, the less likely the user was lost. An average lostness score was calculated for each participant across the three trials at each delay level

Subjective measures. Two 5-point Likert scales assessing frustration and task difficulty were administered after completing each information retrieval task. The rating scale was anchored, such that five was rated as very frustrating or very difficult. The task difficulty measure was used as a manipulation check to ensure the level of complexity of the tasks.

RESULTS

Lostness measure. An average lostness score was calculated for each participant across the three trials at each delay level. A randomized block ANOVA for delay length (1 s, 30 s, 60 s) was conducted and no main effect of delay was observed, $F(2, 26) = 1.47$, $p = .249$; $\eta^2 = .10$; $1 - \beta = .29$. Means for lostness were: .50 (SD = .17) for 1 second, .50 (SD = .15) for 30 seconds, and .58 (SD = .14) for 60 seconds. No main effect of trial was found across the delay conditions, $p > .05$.

Frustration measure. An average frustration measure was calculated across the three trials at each delay level. A completely randomized ANOVA for delay on frustration revealed a significant main effect of delay, $F(2, 26) = 4.15$, $p = .027$; $\eta^2 = .24$; $1 - \beta = .68$. Means for frustration were: 1.9 (SD = .66) for 1 second, 2.83 (SD = 1.00), and 2.76 (SD = 1.20) for 60 seconds. Paired sample t-tests with Bonferroni adjustments were conducted and significant mean differences were observed for frustration between 1 and 30 second delays, $t(13) = -3.05$, $p = .009$, and between 1 and 60 second delay conditions, $t(13) = -2.49$, $p = .027$. No differences were found for frustration between 30 and 60 second delays, $p > .05$.

Task difficulty measure. An average task difficulty measure was calculated across the three trials at each delay level and a randomized block ANOVA for delay length was conducted. Means for task difficulty were: 2.26 (SD = .59) for 1 second, 3.02 (SD = 1.17) for 30 seconds, and 2.76 (SD = 1.01) for 60 seconds. No effect of delay was observed, $p > .05$. A randomized block ANOVA for task revealed no differences across the nine information retrieval tasks, $p > .05$.

Task success. Frequencies for task success are presented in Table 1.

Table 1
Frequency of Task Success across Trials for Each Delay Level

	1 s delay	30 s delay	60 s delay
Success	30	26	27
Unsuccessful	12	14	15
Total	42	40*	42

* 2 missing values in 30 second delay condition

DISCUSSION

The hypotheses were: 1) 30 and 60 second delay conditions would increase lostness on the web site and decrease task success more than 1 second delays, and 2) frustration would be higher on 30 and 60 second delays than 1 second delay conditions.

In the experiment, no effect of delay was observed for lostness. The longer delays did not affect performance on the web site, although the effect size is quite small ($\eta^2 = .10$) and only 14 participants completed the experiment. The frequency of task success was slightly higher for the 1 second delay condition than the 30 and 60 second delays. More participants quit the tasks altogether in the 60 second delay condition than the 1 or 30 second delay conditions. The cash incentive was incorporated into the experiment to increase experimental realism and motivation, but lostness was still not affected by the longer delays. One problem may be that the delay times were fixed, so participants expected a delay and used the time to incorporate different strategies to find task information in the experimental setting than they would typically use. Participants were observed rereading tasks and continuing to search the web page during delays, whereas in reality they may have gone to another site, worked on something else while waiting, or quit searching altogether. These strategies may have improved performance on the web site, so the results may not generalize due to the limitations of the experimental setting.

Frustration was affected by longer download delays in this experiment. Users were more frustrated by the longer delays, with differences occurring between the 1 and 30 second delays and 1 and 60 second delay conditions. A 20 second delay was used in a pretesting experiment, and this delay length was not long enough to frustrate users, which is contradictory to Nielson (1996), who suggested that users would not tolerate download times that exceed 10 seconds. Although, tolerance for delays could also be related to task demands, web page content, user goals or attributes of the population of users sampled, such as computer experience level and download speed they typically experience.

CONCLUSION

When examining response times for web sites, of course faster is always better. Even though lostness was not affected by longer delays, frustration was significantly increased. Therefore, it is imperative to design web sites with download time as an important consideration. Web site performance is an integral part of a company's image and can impact e-commerce sales. If users are frustrated by long download times while visiting a web site, they are likely to leave and go to a competitor's web site. The 1999 projected losses in e-commerce sales, related to unacceptable download

delays, are up to $4.35 million (Wilson, 1999). Limiting the use of large and unnecessary graphics and limiting multimedia effects can decrease download delays for web sites.

There are some problems associated with evaluating the effects of download delays on a web site in an experimental setting. First, the participants are not necessarily interested in the information or the tasks, since the experimenter has selected the tasks and content of the web site. The results may be quite different if the participants were searching for information of interest to them. The users' goals and context of the situation should also be considered. If someone is just browsing the Internet for entertainment, they may be more tolerant of download delays than if they have to find the information to complete a project to meet a deadline. Another important aspect that influences tolerance for time delays is whether the information was worth the wait or considered valuable.

Future research with web site response times should examine the effect of variable delays and focus on web sites of interest to the user. The results may potentially have more external validity, since the delays would be more realistic than the fixed delays used in this study. The delays would also be less predictable, so participants would not be able to set consistent expectations of wait time. In addition, participant's motivation may be increased if the web site content was of interest to them.

REFERENCES

Graphic, Visualization, & Usability Center (GVU) (1998). GVU's 9th WWW user survey. http://www.gvu.gatech.edu.

Kirk, R. (1995). Experimental design: Procedures for the behavioral sciences. Pacific Grove, CA: Brooks/Cole Publishing Company.

Nielson, J. (1996) Top ten mistakes in web design. Alertbox, http://www.useit.com.

Nielson, J. (1997). The need for speed. Alertbox, http://www.useit.com.

Ramsay, J., Barbesi, A., & Preece, J. (1998). A psychological investigation of long retrieval times on the World Wide Web. Interacting with Computers, 10, 77-86.

Sears, A., Jacko, J., & Borella, M. (1997) Internet delay effects: How users perceive quality, organization, and ease of use information. CHI 1997 Proceedings, http://www.acm.org/sigchi/chi97.

Selvidge, P. (1999, May). Reservations about the usability of airline web sites: A comparison of seven airline web sites. Poster session presented at the annual meeting of CHI 1999, Pittsburgh, PA.

Smith, P. (1996). Towards a practical measure of hypertext usability. Interacting with Computer, 8 (4), 365-381.

Wilson, T. (1999, August). Internet Week. http://www.techweb.com/se/directlink.cgi?INW19990802S0002