Web Site Delays: How Tolerant are Users?

Dennis F. Galletta (contact author)
Associate Professor of Business Administration
Katz Graduate School of Business
University of Pittsburgh
Pittsburgh, PA 15260
Phone 412-648-1699
Fax 412-648-1693
galletta@katz.pitt.edu

Raymond Henry
PhD Candidate
Katz Graduate School of Business
University of Pittsburgh
Pittsburgh, PA 15260
Phone 412-648-1647
Fax 412-648-1693
rhenry@katz.pitt.edu

Scott McCoy
Assistant Professor of Management Information Systems
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609-2280
Phone (508) 831-5266
Fax (508) 831-5720
smccoy@wpi.edu

Peter Polak
Assistant Professor of Computer Information Systems
University of Miami
Coral Gables, FL
Phone 305-284-8044
Fax 305-284-5161
ppolak@miami.edu

March 28, 2003

Submitted for Consideration in *Journal of AIS*
This study was presented at the first AIS SIG-HCI Workshop, Barcelona, December 2002

Web Site Delays: How Tolerant are Users?

Abstract

Web page loading speed continues to vex users, even as broadband adoption continues to increase. Several studies have addressed delays both in the context of web sites as well as interactive corporate systems, and a wide range of "rules of thumb" have been recommended. Some studies conclude that response times should be allowed to grow to no greater than 2 seconds while other studies caution on delays of 12 seconds or more. One of the strongest conclusions had been that complex tasks seemed to allow longer response times. This study examined delay times of 0, 2, 4, 6, 8, 10, and 12 seconds using 196 undergraduate students in an experiment. The subjects were randomly assigned a constant delay time and were asked to complete 9 search tasks, exploring a familiar and an unfamiliar site. Plots of the dependent variables performance, attitudes, and behavioral intentions, along those delays, suggested the use of non-linear regression, and the explained variance was in the neighborhood of 2%, 5%, and 7%, respectively. Focusing only on the familiar site, explained variance in attitudes and behavioral intentions grew to about 16%. A sensitivity analysis implies that decreases in performance and behavioral intentions begin to flatten when the delays extend to 4 seconds or longer, and attitudes flatten when the delays extend to 8 seconds or longer. Future research should include other factors such as expectations, variability, and feedback, and other outcomes such as actual purchasing behavior, to more fully understand the effects of delays in today's web environment.

Web Site Delays: How Tolerant are Users?

1. Introduction

One of the most often-discussed objections of the web experience is the frequent delay encountered while browsing. A delay is encountered when a user clicks on a hyperlink and nothing seems to happen for several seconds. Delays have many causes (Nah, 2002), and it is not likely that they will disappear any time soon. Unfortunately, an early study found that waiting time is the most objectionable deficiency of the medium (Lightner & Bose, 1996).

Contrary to expectations, faster connection technology on the Internet has not eliminated the response time problem. Additionally further increases in the speed of technology might have only marginal bearing on alleviating delay problems. Four years ago, Nielsen (1999b) forecasted that low-end users will have to endure unacceptable download times until 2008. The slow improvements are due to (1) the exponential growth of the number of web users in recent years, causing global waiting lines at popular sites, and (2) the unavailability of a low-cost solution to large numbers of web users, tying many users to their modems for the foreseeable future.

The most aggressive improvements in bandwidth will not solve the bottleneck problem; on the contrary, the delays might worsen. An increased number of users operating at high speeds can make more page requests per unit of time, which could put a strain on servers. Web sites might need a much larger number of servers to provide adequate service. For example, with over 10,000 servers (Searchenginewatch, 2002), Google has fast response. However, even with that extreme capacity, which cannot be duplicated by many sites, there are sometimes delays from time to time during heavy use.

Because these delays will be with us for some time, it is important to explore the relationships between delays and performance, attitudes, and behavioral intentions of visitors. There are unfortunately few answers to delay-related questions. How tolerant are users to delay? What is a tolerable delay? What are the effects of longer and longer delays?

This paper will explore the relationship between delay time and user performance, attitudes, and behavioral intentions. The next section explores previous literature in this area. The balance of the paper describes a laboratory study that was conducted to answer the general research questions.

2. Previous Literature

For many years, researchers have investigated a variety of issues related to delays and their effects on users. The Human-Computer Interaction (HCI) area has provided many studies that focus on the examination of delay situations created by computer systems. With the explosion of the World Wide Web, slow computer response times suddenly became salient to the general public. Notwithstanding the numerous advantages that the web provided to computer users, the slow speed of interaction emerged as a serious hindering factor to web usage. Nielsen's May 1999 survey of the official sites of large corporations revealed that of the top ten design mistakes, slow download times took first place with 84% of the sites judged too slow (Nielsen, 1999c, 1999d). Also, the average download time of a page was found to be 19 seconds over an ISDN connection. According to Nielsen's study, users are not sympathetic of waiting for web content; slow response times lead to lower levels of trust toward the web site owner and reliably result in a loss of traffic to the site. Indeed, research has shown that users will leave a site if page load times become too long (Rose et al., 2001).

2.1. Theoretical Foundations

The human-computer interaction literature dates back to the late 1960s when early experiences with time-sharing computer systems included significant waits. The delays were caused by uneven system loads and varied greatly with the number of concurrent users.

One of the most widely-cited papers addressing the problems with delay experienced by people while working with a computer was a conceptual piece by Miller (1968), which followed other related works such as Simon (1966), Newman (1966), and Sackman (1967). Miller proposed a set of guidelines with respect to maximum allowable delays that apply to various end-user tasks. The recommendations were guided by an analogy, in which a human interacting with a computer can be compared to a conversation between two people. Miller suggested a maximum delay of 2 seconds for the interaction not to lose its conversational nature, with about 0.5 seconds being the value resulting in highest conversational flow.

The recommendation to keep delays under 2 seconds has been heavily cited, and has been upheld as a “gold standard” of web design well into the 1990’s (Nielsen, 1999a). In fact, references to the 2-second rule in web page design are so pervasive and the rule’s face validity are so numerous and well-established that its empirical origins are rarely questioned.

The conceivable areas of influence that delays can have on users' interactions with computers have been formulated by Carbonell *et al.* (1968). The authors suggested that delay length can affect users' "behavior, the amount of work they are able to accomplish, and their degree of satisfaction with a time-sharing system" (p.137). Thus, the foundations were in place for establishing dependent constructs, such as attitudes, behavioral intentions, and performance.

2.1.1. Attitudes

Attitudes about information systems have long been studied in experiments, surveys, and field studies. Attitude measures have been used as surrogates for success at nearly all levels of granularity and have been assessed at the individual, group, and organizational levels of analysis. The literature has provided some understanding of how delay and attitudes, in the areas of outcomes, attributions, and other factors, interacting with delay, that affect attitudes. Table 1 provides a list of studies in the area of attitudes with their major findings.

Each of the more detailed areas shown in the Table will be covered, in turn, in the discussion below.

Table 1: Studies of Attitudes and Delay

Study	Major Findings
Outcomes of Delay	
• Ramsay et al., 1998	Faster pages were more interesting and easier to scan
• Hoxmeier & DiCesare (2000)	In a study of delay of 3, 6, 9, and 12 seconds, satisfaction was constant through the 9 second delay condition and dropped for the 12 second delay condition.
Attributions of Delay	
• Jacko et al., 2000	Users often attribute delay to the excessive use of graphics on web pages. Slow sites with graphics led to perceptions of lower information quality and site organization than for text pages.
Estimates of Delay	
• Weinberg, 2000	When users expected slower response time, their estimates of actual response time were longer. Perceptions of quality were not affected by expectations.
Delay and Other Factors	
• Authors, 2003	Attitudinal effects of delay depended on a site's breadth and a user's familiarity with the site. Delay was most tolerable in broad, familiar sites.
• Authors, 2002	Attitudinal effects of delay were found to be alleviated, in part, by feedback during long delays. Variability of the delay did not seem to affect attitudes.

Outcomes of Delay

In a study of latencies ranging from 2 seconds to 2 minutes, Ramsay et al. (1998) found that faster pages are seen as more interesting than slower pages. The study also found that pages loading more quickly were seen as easier to scan than their slower-loading counterparts.

Hoxmeier & DiCesare (2000) employed a simulated web environment and subjects were engaged in an information retrieval search task using download delays of 0, 3, 6, 9, and 12 seconds. The results supported a significant relationship between satisfaction and delay, with satisfaction being highest in the 0-second delay condition.

Satisfaction remained fairly constant throughout the 3 to 9-second range with a noticeable drop in the 12-second delay condition.

Attributions of Delay

Jacko *et al.* (2000) provided evidence that users attribute slow download speeds to an excessive amount of graphics on web pages. When speeds were slow, users perceived lower information quality and lower site organization for sites with graphical pages than for sites with text pages. Apparently, users blamed the slow speed on the excessive use of graphics, which was considered to be a design choice made by the site provider. Because the design was within the provider's control, higher negative attributions were made than in the case of slow text pages where it was perceived that the designer could do little to improve the download speed. Due to the possible interaction among various design elements and download speed, it is necessary to carefully control factors that could influence users' responses to dependent measures.

Estimates of Delay

The role of perceptions on user estimation of download time and satisfaction with a web site was studied by Weinberg (2000). Web site visitors were presented with a message informing them about the subsequent wait. The two treatment conditions consisted of a message stating the wait was 5 or 10 seconds. Under both circumstances, the actual wait duration was 7.5 seconds. The results showed that users in the 5-second message condition reported significantly lower estimates of the waiting time (5.62 s) than the users in the 10-second message condition (8.66 s). However, no significant differences in perceived web site quality were found between the treatment conditions.

Delay and Other Factors

Attitudes were also examined in a comprehensive study of web design factors affecting user attitudes, intentions, and performance conducted by Authors (2003). The effects of download speed with which pages loaded, structural depth of the web site, and familiarity level of the web content were varied. The delay construct assumed two values: an instant response with no delay and a response with a delay of 8 seconds. Subjects' attitudes were more favorable with faster, broader, and more familiar sites. A 3-way interaction among the factors showed that designers can make up for long delays with a broader structure and organizing the site along familiar categories.

In a follow-up investigation, Authors (2002) extended the previous study by examining the effects of delay length, delay variability, and feedback given to users while waiting for page loads in a simulated web browsing session. The delay was manipulated to be, on average, 2 or 10 seconds from a user's click on a link until a complete page load, with either no variability in time intervals between viewing successive Web pages (always 2 seconds or 10 seconds) or 60% variability in both directions from the respective means. A third factor was feedback (providing or withholding process feedback through partially and increasingly displaying the page content as it loaded). Attitudes were most strongly predicted by the delay itself, although providing feedback during long delays helped to alleviate the negative effects of delays to some extent.

2.1.2. Behavioral Intentions

Website quality research has been a focus of multiple academic and practitioner publications. A Web user's behavioral intentions (to return to a site) can serve as an excellent summary variable that indicates content and/or design success, leading to increased popularity or revenue or both. Even in a site that does not involve sales, increased popularity can provide the basis for higher advertising rates, leading to more revenue. Table 2 provides a summary of studies that considered delays and behavioral intentions.

Table 2: Studies of Behavioral Intentions and Delay

Study	Major Findings
Outcomes of Delay	
• Ranganathan & Ganaphy, 2002	Slower pages cause users to seek alternative sites
• Authors, 2003	Behavioral intentions were more favorable with faster, broader, and sites with more familiar structure.
• Hoxmeier & DiCesare, 2002	Intentions to revisit a site decreased significantly as delays increased from 9 to 12 seconds.
• Rose et al., 2001	Users showed a predisposition to abort loading of an e-retailer's web page as delays extended.

In general, the findings on behavioral intentions are similar to those on attitudes. Slow response time can frustrate consumers causing them to seek alternative sites (Ranganathan & Ganapathy, 2002). Authors (2003) found in their study of speed, familiarity, and depth that subjects' intentions were more favorable with faster, broader, and more familiar sites. Hoxmeier & DiCesare's (2000) study of the effects of delays of 0, 3, 6, 9, and 12 seconds revealed that intentions of system reuse decreased significantly in the 12-second category. Rose *et al.* (2001) studied web page delays of 0, 15, 30, 45, 60, and 75 seconds, and found that delay had a significant impact on users' intentions to abort loading of an e-retailer web page.

2.1.3. Performance

The HCI literature has focused substantial attention on performance as "the" central dependent variable for users of technology. System design factors in both hardware and software arenas have been assessed with user-system performance as the "ultimate concern" (Card *et al.*, 1983, p. 404). Performance has been studied for decades, predating the world wide web. Table 3 provides a summary of studies relevant to delay and performance.

Table 3: Studies of Performance and Delay

Study	Major Findings
Preweb Performance Outcomes of Delay	
• Yntema, 1968	As delays increased, users became more efficient, using fewer steps to solve a problem.
• Goodman & Spence, 1978	Time to solution increased by 50% when system response time doubled.
• Thadhani, 1981	User productivity and user response time improved by faster system responses.
• Bergman et al., 1981	Faster system response time led to less efficient strategies
• Butler, 1983	Faster system response time did not affect typing time or correctness of entries for either a simple or more complex task. Faster system response time led to faster user response time only for the simple task.
• Dannenbring, 1984	Faster system response time led users to commit more errors.
Web Delays and Performance Outcomes	
• Authors, 2003	Subjects were able to complete more tasks when delays were short, the site was constructed using familiar terminology, and the site was broad rather than deep. All three factors interacted (both 2-way and 3-way interactions were significant). Performance outcomes were stronger than outcomes of attitudes or behavioral intentions.
• Authors, 2002	Delays predicted user performance more strongly than page loading feedback or variability in loading speed.
• Nah & Kim, 2000	Subjects waited a very long time before aborting the task only for the first non-loading link. Subsequent wait times were about a fifth as long.

Pre-Web Performance Outcomes and Delay

HCI research commenced with studies of user strategies in problem solving situations and ways in which users adapt to various delay conditions. Results reported by Yntema (1968) showed that as users had to work with longer delays, their work strategy indeed changed to accommodate the response characteristics of the system. Users took more total time to solve the problem but they achieved a solution in fewer steps. As the interactions with the system became more costly in terms of time, the users became more careful with system usage and used fewer computer resources.

A study by Goodman & Spence (1978) investigated the effects of increasing system response time (SRT) on users' performance. The authors found that time to solution increased by 50% when the SRT doubled from 0.7 seconds to 1.5 seconds. A further increase of SRT to 3.2 seconds resulted in a decrease in users' performance by an additional 50%, and caused users to complain about the intrusive nature of the system's operation.

In an empirical investigation of user response time and interactive user productivity, Thadhani (1981) found a significant correlation between performance variables and system delay. The data showed that user productivity and user response time were significantly higher in the 0.25-second to 1.0 second response range than they were in the response range greater than 1.0 second.

However, Bergman *et al.* (1981) reported results that were different from those obtained in previous studies. The authors experimented with system response times of 0 and 10 seconds and the results showed no positive effects of short delay on performance measures that were similar to those used in prior research. Total time and total number of trials to reach a solution, total and per trial user response time, and number of trials per minute were among the performance characteristics studied. The data showed that more trials were needed to solve a problem in the immediate response condition than in the 10-second delay condition, which supported the adaptive strategy of users in situations when delays become costly, as previously observed by Yntema (1968), described earlier. That is, as delays become longer, users reduce the number of steps taken.

Butler (1983) conducted two experiments with experienced users performing tasks on cognitively different levels. One task involved simple data entry, and the other was a more demanding information retrieval and record modification task. System response time under investigation assumed a wide range of possible values: 2, 4, 8, 16, and 32 seconds. Again, performance variables served as dependent measures and consisted of typing time, the percentage of incorrect entries, and user response time. Butler's analysis indicated that the mean typing time and the percentage of incorrect entries were not significantly affected by the amount of delay for either task. The only observed relationship was between delay and user response time in the data entry task. Butler's conclusion that "the degradation in user performance seen when average response time is increased appears to be very similar for tasks that are cognitively very different," (p.62) seems to remain uncontested, as we could find no other experiment with cognitively different levels of experimental task measured at various response times.

Dannenbring (1984) conducted an experiment in which beginners and experienced programmers alike were engaged in a quite complex task of debugging a short computer program. User performance and user satisfaction were studied under conditions of 0-second, 5-second, and 10-second system delays, and instructions to work as quickly as possible were given. The only performance variable found significantly related to system delay was the number of corrections of erroneous entries, with fewer characters or lines deleted as delay increased. The findings suggested that faster system response times accommodate the subjects to make more errors, a result consistent with

Bergman *et al.* (1981). Interestingly, user satisfaction with the system, perceived difficulty of the task at hand, total time to solution and other performance measures failed to show a general relationship with computer response time.

Web Delays and Performance Outcomes

Performance was another dependent variable in the study by Authors (2003) discussed above. Subjects were able to complete more tasks with faster, broader, and more familiar sites. In addition, of all of the dependent variables, performance measures were influenced most strongly by interactions among the factors. Performance was significantly affected by all of the two- and three-way interactions between site depth, familiarity, and speed.

In an investigation by Authors (2002), where delay length, delay variability, and feedback given were manipulated, the delay was the most significant predictor of performance on a search task. Although providing feedback (gradually loading graphics) to subjects during long delays helped to alleviate the negative effects of delays to some extent, the absolute length of delay was critical to the users' performance.

Another feedback study was performed by Nah and Kim (2000). In this creative study, ten links were provided to subjects, and only seven worked. Subjects who were provided feedback (a status bar) waited on average 38 seconds after clicking on the first non-working link before giving up, while those without feedback waited only 13 seconds for the first-non-loading link. Their patience did not last long, however, for in subsequent attempts, the subjects only waited 3 seconds without feedback and 7 seconds with feedback. Patience therefore does not seem to last forever.

2.2. Summary

In conclusion, issues related to delays on the Web have recently attracted the attention of researchers. Laboratory experimentation appears to be the primary method of investigation, as it enables strong control over the environmental settings. The results indicate that users prefer Web sites with short delays over sites with long delays, as one might expect. Also, fast sites seem to encourage exploration and decrease the penalty for making errors.

Although these findings are quite consistent across different studies, there is neither comparability among the findings nor data available representing enough intervals to understand the user's tolerance for delay. Also, each previous study included a limited variety of variables, so it is difficult to tie performance, attitudes, and behavioral intentions to specific levels of delay.

This study therefore examines the effects of increasing delay on user performance, attitudes, and behavior intentions, using a laboratory study.

3. Method

Speed was manipulated by using a Javascript program on each page to provide a randomly-assigned constant delay of 0, 2, 4, 6, 8, 10, or 12-second delay per page. Considering a delay as "long" is quite subjective and it was therefore difficult to make the choice considering the wide range used in previous experiments. We wanted to have enough settings to include at least two values above a commonly-discussed maximum of eight seconds (Hoxmeier & DiCesare, 2000; Ramsay *et al*., 1998; Zona, 1999; Shneiderman, 1998; Kuhmann, 1989). Also, our simple site did not seem to contain enough graphical content to justify a longer delay, which could lead to unpredictable attributions (Sears & Jacko, 2000).

3.1.1. Materials

Two artificial Web sites were taken from a previous study (Authors, 2003), and included "Pete's General Store" as the "familiar" site and "A.C.T. Systems" as the "unfamiliar" site. The familiar site contained groceries and/or home products, arranged in easily recognizable categories such as "Health Care Products" and "Food Products." On the other hand, the unfamiliar site contained fictitious products arranged into categories that provided no clue as to their meaning, such as "Novo Products," and included completely fictitious software products and computer accessories. Both sites included brief product descriptions, prices, and images, and the search tasks led to pages in precisely the same position on each site to provide at least some measure of comparable difficulty in the two sites.

Users were asked to search in both sites. To prevent confounding order effects, for half the subjects, the familiar site was presented first and for the other half, the unfamiliar site was presented first.

3.1.2. Subjects

A total of 196 subjects volunteered to participate in the study. Subjects were solicited from a population of upper-level undergraduate business students at a large university in the northeast United States. All of the participating faculty members offered students extra credit for participating, so nearly 100% of the subjects volunteered.

Students were considered to be appropriate subjects for this study because the experimental task does not focus on contextual factors and decision-making situations, and we expect most people to react in a similar way to increases in web delays. That is, we focused on what we would expect to be “invariant” (Simon, 1990) across individuals. Further, Voich (1995) found values and beliefs of students to be representative of individuals in a variety of occupations.

3.1.3. Measures

Performance was measured by totaling subjects’ scores on nine search tasks that were developed for this study. One point was earned per correct fill-in-the-blank answer. Each task required subjects to browse the site to find details about products such as packaging or pricing. Although short, dichotomous instruments are not normally subjected to reliability analysis, the Kuder-Richardson-20 statistic, or KR-20 test (analogous to alpha for dichotomous items), was quite high in this study (.90).

Attitudes about the sites were measured by summing the responses from seven 9-point Likert-type questions that were adapted from Part 3 of the long form of the QUIS (Questionnaire for User Interaction Satisfaction) (Shneiderman, 1998, p. 136), which has been tested for reliability and validity in previous research (Chin *et al.*, 1988). We dropped one item to maximize Cronbach’s alpha, which was .86.

Behavioral intentions were measured using the sum of two, related, original 7-point Likert-type questions: how readily the subject would visit the site again and how likely he or she would recommend that others visit the site. The alpha score for this very short instrument was also extremely high, at .94.

3.1.4. Procedure

The three Web sites (practice, familiar, and unfamiliar task) were written on CDs to precisely control the browser’s response time. All subjects used identical PCs and 17”

XGA screens in a campus lab containing 46 machines. After signing an "informed consent" form, subjects were told that participation was voluntary and that they could leave the study at any time. A randomly-assigned code found on their packet was entered on the screen, activating the proper delay treatment. The experimenters ascertained that each subject entered the correct number from the packet.

A small practice site provided a chance for subjects to become familiar with the search task, and then the two main sites were assigned. In each main site, the 9 tasks were then undertaken. After completing the tasks as best they could, subjects were instructed to close the browser window and complete the questions addressing attitudes and behavioral intentions. The entire exercise took, on average, an hour to complete.

4. Results

In graphical form, the basic results are shown in Figure 1 as error bars defining means and 95% confidence intervals at each time setting. In general, the results illustrate, as expected, declining performance, attitudes, and behavioral intentions as delays increased. One exception is that there appears to be a performance increase at the 8 and 12 second levels. This result could indicate changes in strategies, found in other studies (Yntema, 1968), although there is not sufficient evidence to strongly support this assertion.

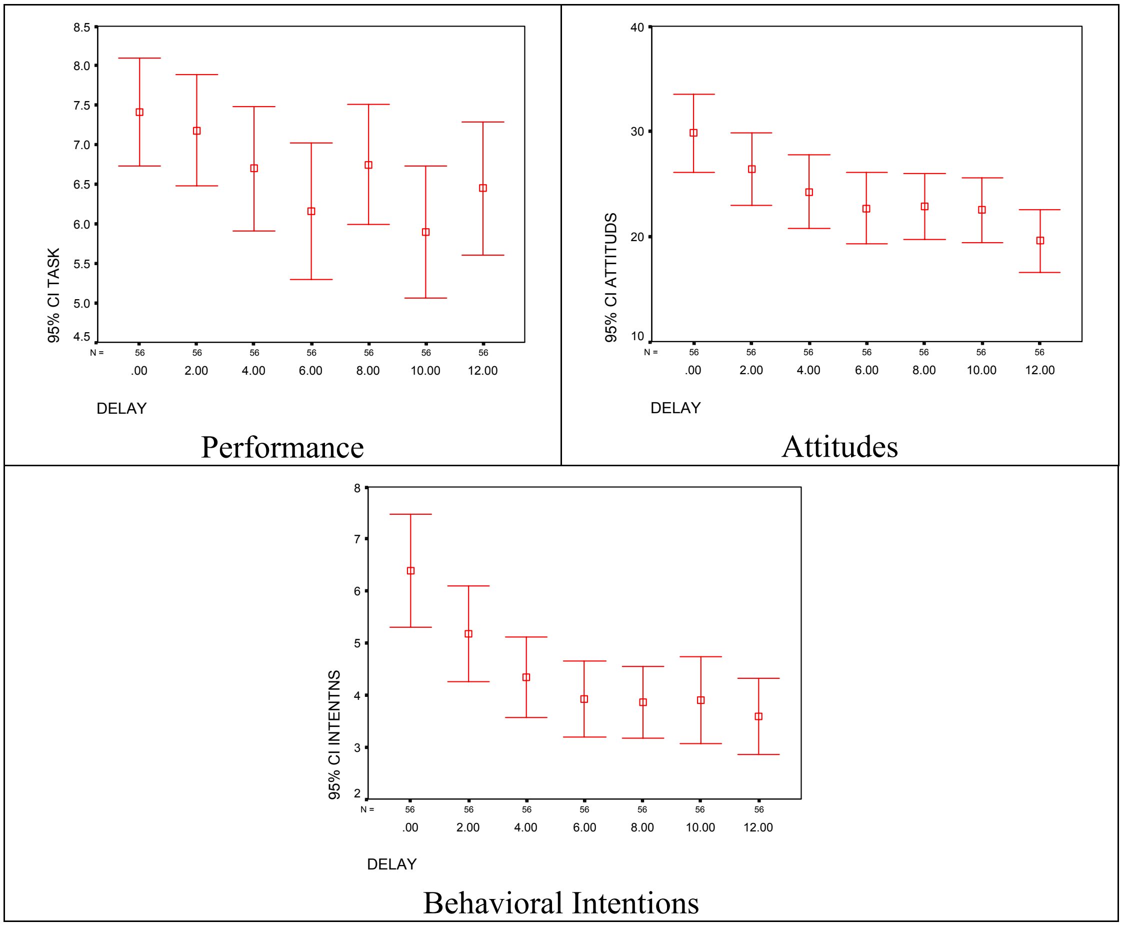

Figure 1: Error Bars for each Dependent Variable

Curvilinear regression served to estimate curves and explained variance for each dependent variable for two reasons. First, the graphs in Figure 1 clearly exhibit a non-linear progression, as would perhaps be expected because predictions for the dependent variables could not become negative as time grows to infinity. Stated another way, the expected and graphed results are likely to be asymptotic with respect to the X axis. Second, Davis & Hantula (2001) stated that response curves should not necessarily be considered linear. Linear regression results will also be presented, largely as a benchmark for comparing the curvilinear results. In all cases a logarithmic function provided nearly the best fit, even if log functions are not necessarily asymptotic to the X axis at larger values of the delay.

Results of regression testing are shown in Table 4. As expected, the curvilinear regression explains more variance than the linear approach, although both approaches provide similar results. The logarithmic and linear regression equations are significant for all of the dependent variables, but the amount of explained variance is relatively modest (explained variance for performance, attitudes, and behavioral intentions are 1.9%, 5.2%, and 7.9%, respectively).

Table 4: Regression Results – All Subjects

Dependent	Method	Rsq	d.f.	F	Sigf	b0	b1
Performance	Logarithmic	0.019	390	7.36	0.007	7.0908	-0.3213
Performance	Linear	0.017	390	6.84	0.009	7.2518	-0.0966
Attitudes	Logarithmic	0.052	390	21.37	0	27.2138	-2.3246
Attitudes	Linear	0.05	390	20.67	0	28.4685	-0.7133
Intentions	Logarithmic	0.079	390	33.67	0	5.4628	-0.73
Intentions	Linear	0.064	390	26.63	0	5.7321	-0.2041

Separate analysis of the familiar and unfamiliar sites revealed that the explained variance in performance and significance of the regression equations improves dramatically.

When data from only the unfamiliar site are used, all regressions remain highly significant, but for performance the equations explain substantially more variance, and for attitudes and behavioral intentions the equations explain substantially less variance. As is shown in Table 5, 6.1% of the variance in task performance is explained in the curvilinear regression for the unfamiliar site only (5.3% in the linear regression), which is about three times the corresponding value in Table 4 for all subjects.

Table 5: Regression Results – Unfamiliar Site Only

Dependent	Method	Rsq	d.f.	F	Sigf	b0	b1
Performance	Logarithmic	0.061	194	12.71	0	5.3582	-0.5598
Performance	Linear	0.053	194	10.76	0.001	5.5952	-0.1614
Attitudes	Logarithmic	0.058	194	11.88	0.001	16.8726	-1.4032
Attitudes	Linear	0.037	194	7.39	0.007	17.1191	-0.3489
Intentions	Logarithmic	0.05	194	10.27	0.002	3.0935	-0.29
Intentions	Linear	0.029	194	5.72	0.018	3.1204	-0.0682

When data from only the familiar site are used, regressions for attitudes and behavioral intentions again remain highly significant, but performance loses significance. Curvilinear equations for attitudes and behavioral intentions explain substantially more variance (15.5% and 17.7%, respectively) in this sub-sample than in the overall sample or in the unfamiliar data only, as shown in Table 6.

Table 6: Regression Results – Familiar Site Only

Dependent	Method	Rsq	d.f.	F	Sigf	b0	b1
Performance	Logarithmic	0.012	194	2.33	0.129	8.8233	-0.0828
Performance	Linear	0.018	194	3.57	0.06	8.9085	-0.0319
Attitudes	Logarithmic	0.155	194	35.54	0	37.5551	-3.246
Attitudes	Linear	0.176	194	41.34	0	39.8179	-1.0778
Intentions	Logarithmic	0.177	194	41.67	0	7.832	-1.1701
Intentions	Linear	0.154	194	35.2	0	8.3439	-0.3399

4.1.1. Sensitivity Analysis

One of our research goals was to learn at which point delays cease to have an effect on the dependent variables. Similar regression analysis was run for subsets of the data in both familiar and unfamiliar cases. At each step, data from the lowest remaining time delay were removed to see if a significant regression equation could be computed.

While eventually this approach would run out of statistical power, this procedure is followed simply as an exploratory step. Further research is needed to provide additional understanding of the nature of the curve as delays become quite lengthy.

After removing the zero-delay subjects, regressions were no longer significant for performance, but remained significant for attitudes and behavioral intentions ($p < .005$). Explained variance was, coincidentally, 2.4% for both measures. After the next step, removing the 2 second subjects, no regressions were significant. Therefore, across both familiar and unfamiliar sites, one might argue that any delay above 2 seconds is detrimental, but that further delays do not provide important reductions in the outcomes we examined.

Analysis was repeated for the separate sites. For the unfamiliar site only, only performance survives the loss of the zero-delay subjects. More specifically, the linear regression ceases to be significant but the curvilinear regression remains significant ($p<.023$) and explains 3.1% of the variation in performance. Therefore, users in an unfamiliar site have significant reductions in attitudes and behavioral intentions with any delay at all, and suffer performance degradation when exceeding 2 seconds in delay.

For the familiar site only, where only attitudes and behavioral intentions are significant to start, dropping the zero-delay subjects results in regressions that continue to be significant and explain variance. Total delay from 2 seconds and above explains 10.6% of the variance in attitudes ($p < .001$) and 6.8% of the variance in intentions ($p<.001$) in the linear regressions. Results were similar for curvilinear regressions, so only the linear results will be discussed.

After dropping the 2-second delay subjects, only the attitude data provided a significant regression equation. Linear and non-linear regression results were nearly identical. Delays from 4 to 12 seconds explained 6.4% of the variance in attitudes ($p = .003$). After dropping the 4-second delay subjects, the attitude regressions remained significant. Total delay from 6 seconds and above explained 3.7% of the variance in attitudes ($p = .041$). No further regressions were significant.

A summary of the results of Sensitivity Analysis can be found in Table 7. The analysis shows that the regressions lost significance suddenly, and performance and behavioral intentions clearly flattened out before the loss of statistical power took over. On the other hand, attitudes showed the clearest persistence of an effect as delays lengthened, but only for the familiar site. In the unfamiliar site, attitudes were slightly more persistent in exhibiting continued decreases beyond the 2-second point, again before the loss of statistical power would have predicted.

On the basis of the analysis, it might be concluded that web designers could "lose" their audience at much lower delays than they might expect; users might be unable to complete their tasks and they might not return if delays reach 4 seconds or more. That is, the "damage is done" at relatively low delay times for task performance and behavioral intentions; after the delay reaches 4 seconds, further delay increases do not further degrade those outcomes. Attitudes, however, continue to degrade until they reach 4 seconds for both sites, and 8 seconds for familiar sites, where they appear to level off or until statistical power is significantly impaired.

Table 7 – Significant Regressions (For each subset of time delays, this table shows for which tasks a significant regression equation could be found)

	0 & up	2 & up	4 & up	6 & up	8 & up	10 & 12
Performance	Both	Unfamiliar Only *				
Attitudes	Both	Both	Familiar Only	Familiar Only		
Behavioral Intentions	Both	Both				

* curvilinear regression results only

5. Discussion and Conclusions

This study examined how increases in delay affect web users. The results indicate that increases in delay clearly relate to decreases in performance, attitudes, and behavioral intentions. These decreases are not necessarily linear, as can be seen from the higher explained variance resulting from the nonlinear regression under several of the conditions tested.

For researchers, one of the most important findings in this research is that relatively small increases in delay can have a profound impact on how users react to web sites. In this study, delays ranged from 0-12 seconds, and significant results were found at the short side of this range. Other studies investigating the effect of speed on users' reactions to web sites have used delays as high as 2 minutes. These results suggest that it is unnecessary to impose such long delays on experimental subjects; users are much more sensitive to delay than would be apparent by examining the ranges used in the previous literature.

For the practitioner, these results also suggest that problems associated with delay will still be present even with the increased use of high-speed connections. Many of the effects of delay will occur at much lower levels than can be addressed simply by changing the connection speed. Fundamentally, if the designer's goal is to encourage the user to "stick with" the task or to return later, then the site delay should be kept below 4 seconds. If the goal is to promote a positive attitude, the site delay should be kept below 8 seconds.

The effects of increases in delay also appear to depend on the familiarity of terminology used in organizing the site. For familiar sites performance does not appear to be affected by changes in delay, at least not at the levels tested in this research. This result may not be surprising. When the content is familiar users are still capable of completing their tasks regardless of the amount of delay they experience. However, increases in delay explain as much as 17% of the variance in user's attitudes and intentions to return to these same familiar sites. For unfamiliar sites, the impacts of changes in delay are also significant, although these changes explain less of the variance. This result suggests that while users are still sensitive to delay it has less of an effect, possibly because of other difficulties associated from navigating in an unfamiliar environment. These results suggest that as users become more familiar with a web site, delays become more salient and play a larger role in formulating attitudes and intentions.

Several limitations must be mentioned. First, in this research all delays were carried out the same way on each page. While the type of delay used is commonly experienced and can be seen in non-laboratory situations, it does not represent all possible manners of delay that might actually occur in internet use. The results, therefore, cannot necessarily be attributed to all types of delays. Future research may explore the effect of different manifestations of delay or variation in these manifestations on performance, attitudes, and intentions. Additionally, this research did not address user expectations or the level of importance that users put on the tasks or information available on the sites. In a laboratory setting it would be difficult to manipulate these factors, however, they could influence users' sensitivity to delays.

This study attempted to bring together factors previously studied separately. A variety of delay levels are coupled with site familiarity in predicting user attitudes, behavior, and performance, three important outcomes examined in previous research. In bringing these factors together, it appears that users are much more impatient than previously thought. If this impatience is indeed task independent, as some, but not all, previous work suggests, subsequent studies might investigate ways in which to reduce the impatience or at least reduce any ill effects resulting from that impatience, such as failure to return to a site or the formation of negative attitudes and/or word of mouth.

Future research should focus on identifying and quantifying factors that interact with the tolerance for delay in an attempt to formulate a more complete model for understanding its antecedents and consequences. Those factors could include some that have already been examined, such as feedback, variability, familiarity, and site depth, and some that have not, such as expectations, involvement with site content, graphics-intensiveness, database-intensiveness, and processing-intensiveness.

Such understanding will allow researchers to provide experimental environments that are reasonable, would allow researchers and practitioners to more realistically assess web design alternatives, and would allow practitioners to be more sensitive to the needs of web users.

6. References

Authors (2002). "The direct and interactive effects of web site delay length, delay variability, and feedback on user attitudes, performance, and intentions," *PhD Dissertation*.

Authors (2003). "The direct and interactive effects of web site speed, familiarity, and breadth on user attitudes, behavioral intentions, and performance," *Working paper*.

Bergman, H., Brinkman, A., & Koelega, H.S. (1981). "System response time and problem solving behavior," *Proceedings of the Human Factors Society – 25th Annual Meeting*, 749-753.

Butler, T.W. (1983) "Computer response time and user performance," *Proceedings of ACM SIGCHI '83 Conference on Human Factors in Computer Systems*, 58-62.

Carbonell, J.R., Elkind, J.I., & Nickerson, R.S. (1968) "On the psychological importance of time in a time-sharing system," *Human Factors*, 10, 135-142.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Chin, J.P., Diehl, V.A. & Norman, K.L. (1988). "Development of an instrument measuring user satisfaction of the human-computer interface," *Proceedings of CHI '88—Human Factors in Computing Systems,* New York, 213-218.

Dannenbring, G.L. (1984). "The effects of computer response time on user performance and satisfaction: A preliminary investigation," *Behavior Research Methods & Instrumentation*, 15 (2), 213-216.

Davis, E.S. & Hantula, D.A. (2001). "The effects of download delay on performance and end-user satisfaction in an Internet tutorial," *Computers in Human Behavior*, 17, 249-268.

Goodman, T.J. & Spence, R. (1978). "The effect of system response time on interactive computer-aided problem solving," presented at *Proc. Siggraph '78 Conf. Association for Computing Machinery*, New York.

Hoxmeier, J.A. & DiCesare, C. (2000) "System response time and user satisfaction: An experimental study of browser-based applications," *Proceedings of the Association of Information Systems Americas Conference*, Long Beach, California.

Jacko, J.A., Sears, A. & Borella, M.S. (2000). "The effect of network delay and media on user perceptions of web resources," *Behaviour and Information Technology*, 19 (6), 427-439.

Kuhmann, W. (1989). “Experimental investigation of stress-inducing properties of system response times,” *Ergonomics*, 32 (3), 271-280.

Lightner, N.J. & Bose, I. (1996). “What is wrong with the world-wide web?: A diagnosis of some problems and prescription of some remedies,” *Ergonomics*, 39 (8), 995-1004.

Miller, R.B. (1968). “Response time in man-computer conversational transactions,” *Proceedings of the Fall Joint Computer Conference*, 267-277.

Nah, F. (2002). “A Study of Web Users’ Waiting Time,” in *Intelligent Support Systems Technology: Knowledge Management,* Vijayan Sugumaran (editor), IRM Press, 2002, pp. 145-152.

Nah, F.H. and Kim, K. (2000) “World Wide Wait,” in *Managing Web-Enabled Technologies in Organizations: A Global Perspective,* Medhi Khosrowpour (editor), Idea Publishing Group, pp. 146-161.

Newman, J.R. (1966). “Extension of human capability through information processing and display systems,” System Development Corp., Santa Monica, California.

Nielsen, J. (1999a). “User interface directions for the web,” *Communications of the ACM*, 42 (1), 65-72.

Nielsen, J. (1999b). “’Top ten mistakes’ revisited three years later,” *Alertbox,* http://www.useit.com/alertbox/990502.html

Nielsen, J. (1999c). “Who commits the ‘Top ten mistakes’ of web design?” *Alertbox,* http://www.useit.com/alertbox/990516.html

Nielsen, J. (1999d). “The Top ten new mistakes of web design,” *Alertbox,* http://www.useit.com/alertbox/990530.html

Nunnally, J. (1978). *Psychometric Theory,* New York: McGraw-Hill Book Company.

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.

Ranganathan, C. & Ganapathy, S. (2002). “Key dimensions of business-to-consumer web sites,” *Information and Management,* 39 (6), 457-465.

Rose, G.M, Lees, J. & Meuter, M. (2001). “A refined view of download time impacts on e-consumer attitudes and patronage intentions toward e-retailers,” *The International Journal on Media Management*, forthcoming.

Sackman, H. (1967). “Experimental investigation of user performance in time-shared computing systems,” System Development Corp., Santa Monica, California.

Searchenginewatch (2002). "The technology behind Google," *SearchDay,* August 12, 2002, available at http://searchenginewatch.com/searchday/02/sd0812-googletech.html.

Sears, A. & Jacko, J.A. (2000). "Understanding the relation between network quality of service and the usability of distributed multimedia documents," *Human-Computer Interaction,* 15 (1), 43-68.

Shneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 3rd edition, Reading, Massachusetts: Addison-Wesley.

Simon, H.A. (1966). "Reflections on time sharing from a user's point of view," *Computer Science Research Review*, 43-51, Carnegie Institute of Technology.

Simon, H.A. (1990). "Invariants of human behavior," *Annual Review of Psychology*, (41), 1-19.

Thadhani, A.J. (1981). "Interactive user productivity," *IBM Systems Journal*, 20 (4), 407-423.

Voich, D. (1995). *Comparative Empirical Analysis of Cultural Values and Perceptions of Political Economy Issues*, Westport, Connecticut: Praeger.

Weinberg, B.D. (2000). "Don't keep your internet customers waiting too long at the (virtual) front door," *Journal of Interactive Marketing*, 14 (1), 30-39.

Yntema, D.B. (1968). "Engineering psychology in man-computer interaction," paper presented at *Annual Convention of the American Psychological Association,* San Francisco, California.

Zona Research Study (1999). "The need for speed," http://www.zonaresearch.com/info/press/99%2Djun30.htm