Abstract

Although its popularity is widespread, the World-Wide Web is well known for one particular drawback: its frequent slow speed when moving from one page to another. This experimental study examined whether speed and two other Web site design variables (site depth and content familiarity) have direct and interactive effects on user performance, attitudes, and behavioral intentions. An experiment was conducted with 160 undergraduate business majors in a completely counterbalanced fully factorial design that exposed them to two Web sites and asked them to browse the sites for 9 pieces of information. Results showed that all three factors have strong direct impacts on user attitudes and performance, leading to behavioral intentions, as predicted. A significant 3-way interaction was found between all three factors on attitudes and performance, and strong 2-way interactions were found between all pairs of factors on performance, also as predicted. The negative attitude and performance effects of slow speed appear to be reduced by breadth and familiarity together, but only performance is enhanced by all possible two-way interactions. Additional research is needed to discover other factors interacting with speed, other effects of slow speed, and under what amounts of delay these effects occur.

Keywords: electronic commerce, response time, web site design, attitudes, performance
ISRL Categories: AA01, AA02, AA05, AI0105, EI0206, FB0403, HA0902, HC0101
When the Wait Isn’t So Bad: The Interacting Effects of Web Site Speed, Familiarity, and Breadth

There is perhaps no innovation in the past fifty years of computing that has attracted as much attention as the Internet. In spite of the large amount of attention paid to business-to-consumer electronic commerce, in the last two years the frantic pace for investment in “dot com” endeavors has subsided and on-line consumer spending has not reached expected levels. Although consumers are expected to spend on average $4.9 billion per month online in early 2003 (Cox, 2003), this amount represents less than 1% of total consumer spending on goods and services, and less than 2% of monthly consumer spending on goods alone (US Bureau of Economic Analysis, 2002).

While relatively few purchases are made using the Web, the majority of users report trying to use the Web to at least collect information for purchase decisions. However, many users have a variety of problems collecting that information. The three most important problems included the failure to find the products they were looking for, confusing or disorganized sites, and speed (GVU, 1998).

For reasons that will be developed in this article, speed is likely to be the core problem that users face in finding information, perhaps posing strategic disadvantage from slow service; losses from download times of 8 seconds and higher are estimated to be $4.35 billion each year (Zona, 1999). Traffic appears to be dependent on page load speed, supporting this claim to some extent; improving page load speed from 8 seconds and higher to between 2 and 5 seconds has doubled the traffic of some sites (Wonnacott, 2000). An analysis of the most popular sites led Nielsen (1999a) to conclude that the enhanced usability of a fast site leads to increased usage.

An observer might indeed conclude that speed is the only concern. However, there are other factors that interact with users’ reactions to delays. This study examines empirically two
potential factors, familiarity and structure, that interact with delay speed in predicting user performance, attitudes, and behavioral intentions.

**Speed: No Silver Bullet**

Speed is considered by some to be “the most commonly experienced problem with the Web” (Pitkow *et al*., 1998), leading many to call it the “World Wide Wait” (Nah & Kim, 2000). Although modal access speed has increased considerably since GVU’s (1998) first survey, problems of delay continue to plague users (Nielsen, 1999b). Many are turning to broadband access to try and solve the problem. Following a year of unprecedented growth, out of 606 million Internet users worldwide (www.nua.com, 2002), an estimated 46 million enjoy a broadband connection (In-Stat/MDR, 2002). Dataquest/Gartner (2002) estimates that only 14% of U.S. households now have broadband access to the Internet.

One should not be tempted, however, to assume that providing broadband to an ever-increasing minority of people will solve speed problems forever. Speed difficulties have a variety of causes (Nah & Kim, 2000; Rose *et al*., 1999; Nielsen, 1997), and the ultimate speed of loading a page is only as fast as the slowest link in the chain from the server itself, to the server’s connection, to the Internet backbone, to the user’s connection, and finally, to the user’s computer and browser. Interestingly, Koblentz (2000) reports that misconfiguration of the server can in some cases triple page loading time.

If they do not suffer from telecommunications congestion at the weakest link in the system, users could find themselves in a global waiting line at the server or victims of design or configuration error. Network traffic growth continues to exceed upgrades in network bandwidth, painting a gloomy picture for the Web of tomorrow (Sears & Jacko, 2000). An increase in the
number of broadband users will require a large number of server upgrades (Connolly, 2000). In the words of Sears and Jacko, “delays will be a concern for many years to come” (p. 49).

**Speed, Familiarity and Structure**

While seemingly unrelated, a case will be made in this paper that familiarity and structure interact with speed in their effects on a user’s Web experiences. Slow loading of a single page or even a small number of pages might not be very noticeable or memorable. However, repeated delays are quite common for two reasons:

- Many sites have “deep” structures that require several clicks to arrive at a target page.
- Many users are unfamiliar with the content and/or structure of the site and become confused or “lost,” choosing the wrong hyperlink (Edwards & Hardman, 1989). Such errors require them to back up and make alternative choices to arrive at the correct target page.

Because of repeated clicks on each site, what someone might consider a short delay for a single page is multiplied by dozens of pages, adding up to a sizable time investment in viewing what might be a blank screen for a great deal of time.\(^1\)

The factors mentioned above lead to the following research question:

*How do speed, familiarity, and site depth interact to influence attitudes, performance, and finally, behavioral intentions in online product information searches?*

This question should be of interest to researchers who wish to understand how those factors independently and interactively form a predictive model of attitudes and performance. It should also be of interest to designers who wish to encourage sales of their products, either directly through on-line purchases or indirectly through customers’ search for product information.

\(^1\) Although using a search engine would allow a user to jump directly to the resource needed, full-text searches are not always available, or when available, are often not very useful. Few people can use such tools effectively (Ziegler, 1996), partly due to the unique syntax and behavior of each search engine (Lohse & Spiller, 1998), partly because searching is a difficult task for most users (Gauch & Smith, 1991; Blair & Marron, 1995), and partly because most search outcomes lack precision (Kaindl *et al*., 1998). In short, the typical search engine interface is “notorious for its lack of usability” (Lee, 1999, p. 38). Sometimes terminology itself yields unintentionally poor search terms, such as searching for the old term “PCMCIA” versus searching for “PC Card,” the new name for the same item.
Because speed is an obvious culprit for web delays for which designers often have little control, it is important to understand ways in which other features of the site might compensate for, or perhaps exacerbate, speed difficulties. While new to the web, finding online information has received attention in menu design and response time literature, described below.

We therefore assert that:

1. Providing fast access to a target Web page is a complex issue that cannot be remedied with a single technology or approach; for the foreseeable future, it appears that users will continue to confront issues of speed.
2. A design choice that increases the depth of a site can greatly magnify the delay.
3. Users will make errors in finding a target page on an unfamiliar site, and those errors will also magnify the delay.

The remainder of this paper outlines theoretical foundations and describes an experimental study where we apply these assertions. Practical and theoretical implications are also offered in light of our results.

**Theoretical Foundations**

This experimental study examines Web site download speed, depth, and familiarity of content and structure, to determine the extent to which these factors have independent and interactive effects on user attitudes, performance, and finally, behavioral intentions.

The research model is shown in Figure 1.

**Speed**

Our understanding of Web page download time can be informed by research on computer response times. The speed of system response has been shown consistently to be a major factor in overall usability both in studies of the World-Wide Web (Nielsen, 1999b) and other large-scale information systems (Shneiderman, 1998; Kuhmann, 1989). Research on delays has focused on the effects of delays and the maximum delays tolerable.
Effects of delay. Prior research has shown an inverse relationship between system response times and user performance (Barber & Lucas, 1983; Butler, 1983) and productivity (Martin & Corl, 1986; Dannenbring, 1983). Longer response times lead to frustration (Doherty & Kelisky, 1979), dissatisfaction (Lee & MacGregor, 1985), frequent feelings of being lost (Sears et al., 2000), and giving up (Nah, 2002). A Web site that manages waiting time effectively can prevent negative affect from carrying over to the site itself, except when delays are longer than expected, occur in unpredictable patterns, or are not accompanied by duration or status information (Dellaert & Kahn, 1999). Unfortunately, on the Web it is nearly impossible to offer any of these remedies reliably.

Tolerable Level of Delay. Over the years, researchers have attempted to identify the maximum amount of delay that would be tolerable to users in a variety of settings. Early studies suggested that the ideal response time was approximately 2 seconds (Miller, 1968), although such short response times might be impractical to use as a general design standard (Shneiderman, 1998), especially on the Internet with its abundant and varied sources for delay described earlier.

What is the maximum delay? Studies support nearly any number beyond the 2-second “ideal.” For example, delays over 38 seconds have led subjects to abandon a task (Nah, 2002). Also, 15 seconds have been described as disruptive (Shneiderman, 1998). Users have also been
found to lose patience when delays reach 12 seconds (Hoxmeier & DiCesare, 2000), to lose interest when delays exceed 10 seconds (Ramsay et al., 1998), and to suffer psychological and performance consequences with delays of 8 seconds (Kuhmann, 1989).

Following up on old literature on response time, Sears and Jacko (2000) present a possible reconciliation of these varied recommendations, suggesting that the amount of frustration on the Web depends on users’ perceptions of the source of the delay. That is, if users see the delay as unavoidable (for example, having large graphics that are the main object of interest), they will have higher tolerance. If the delay, however, is due to unnecessary graphical content accompanying simple text that is of interest, then their impatience will likely be greater.

A recent study by Galletta et al. (2002) imposed delays of 0, 2, 4, 6, 8, 10, and 12 seconds on 196 web-browsing subjects to determine effects of longer times on performance, attitudes, and behavioral intentions. The results revealed significant effects on all three dependent variables as delays increased. Interestingly, by the time delays exceeded 6 seconds, further reductions in performance, attitudes, and behavioral intentions seemed to lose significance.

It is striking that such relatively short delays in prior research have been shown to produce significant results in both user attitudes and performance. Users of the World Wide Web often must deal with much longer wait times (Ramsay et al., 1998). This indicates the importance of understanding the consequences of delays and ways of mitigating those delays on the Internet. Therefore, we propose the following hypotheses:

\[ H1a: \text{A fast site will result in more favorable user attitudes about the Web site than will a slow site.} \]
\[ H1b: \text{A fast site will result in higher user performance than will a slow site.} \]

**Depth**

Especially when response times are long, users can react unfavorably when several selec-
tions must be made in order to complete a task (Shneiderman, 1998). Many alternative Web site
designs can be selected, with varying effects on the number of selections that must be made.
Given a fixed number of end nodes in a Web site, a designer can elect to use a “broad” strategy,
increasing the number of hyperlinks on each page while decreasing the number of clicks and
page loads, or a “deep” strategy with fewer links per page but more hierarchical levels. The
tradeoff between breadth and depth has been widely studied in menu design for information ac-
quisition (Norman & Chin, 1988; Parkinson et al., 1988; Sisson et al., 1986; Landauer &
Nachbar, 1985; Tullis, 1985; Kiger, 1984; Snowberry et al., 1983; Miller, 1981) and Hypermedia
(Edwards & Hardman, 1989).

Much of the work concerning the tradeoff between breadth and depth has focused on de-
termining the optimal number of items per level (Larson & Czerwinski, 1998). While there has
been some variation in research findings, most of the studies listed above have determined that
breadth is preferred to depth (Jacko et al., 1995). The theory of task complexity (Frese, 1987)
can be used to explain this preference (Jacko & Salvendy, 1996): A deeper structure increases
the number of decisions that must be made, which in turn increases complexity, user response
time and the number of errors. This explanation is consistent with Shneiderman’s (1998) obser-
vation: “when the depth goes to four or five (levels), there is a good chance of users becoming
lost or disoriented” (p. 249).

Although increasing the breadth of a site requires fewer clicks by the user to reach the de-
sired page, there might be a tradeoff—providing too many links on a single page can clutter the
page, making it difficult to read or comprehend (Larson & Czerwinski, 1997). This potential for
overcompensation highlights the fact that overall performance is a combination of system re-
response time and the time it takes users to scan the selections and make a decision (Landauer &
Nachbar, 1985). Each decision leads to another page load delay and need for another decision (Zaphiris & Mtei, 1997), so it is not surprising a “broad” structure is most often preferred, especially because most laboratory studies consider 8-9 items per menu to be a “broad” structure and 2-3 items per menu to be a “deep” structure.

In a laboratory setting, menu depth and Web sites are often much smaller than in practice, yet effects are found to be strong. There are several plausible explanations for the strength of the effects found in the laboratory:

1. Users seldom need to experience an entire Web site. They most often attempt to find a particular item, cutting significant portions out of their search with each choice. Thus, browsing is often constrained to a limited area.²
2. Experimental sites created for laboratory studies are highly controlled for the factors of interest, isolating their effects.
3. It is likely that subjects have more intrinsic interest in either a task that is job-related or Web sites they choose to visit, therefore, any design factors that facilitate or hinder task completion are magnified as subjects attempt to finish the task and return to more intrinsically interesting activities.

To summarize, an ideal, universal depth or breadth of a Web site cannot be determined, although most studies have found that users prefer, and perform best with, a “broad” structure with 8-9 items per menu in interactive systems. Therefore, we propose the following hypotheses regarding the depth of web sites:

- **H2a**: A broad Web site will result in more favorable user attitudes about the site than will a deep Web site.
- **H2b**: A broad Web site will result in higher user performance than will a deep Web site.

**Familiarity**

Familiarity with Web site content has been shown to be an important determinant of shopping effectiveness, but familiarity is difficult to control because of the widespread heteroge-

---
² For example, a site with 6 levels of 9 nodes each has 66,430 pages. Making three menu choices, such as region/language, product line, and general item category would reduce the number of pages in the remaining structure to 91. Therefore, an overwhelming majority of the site would never be seen by the typical user after a few high-level choices are made.
neity of users (Chau et al., 2000). The ability to navigate through a Web site (or any hierarchical system) is influenced by a user’s understanding of not only content but also the structure of the site. Content represents the material presented on the pages, and structure represents the categories used to group the branches.

We therefore define familiarity as the ability of users to discern, from the terminology presented on the site, their next move to reach their goal. Because links are very often words that have meaning specific to the content of the site, familiarity of content and structure are expected to be highly interrelated constructs. Once a user has visited a site several times, and the idiosyncrasies are learned, the ability to discern the location of the goal would increase dramatically.

Pages higher on the hierarchy often depict broad identifiers for which the user must make inclusion judgments in order to reach the appropriate page at lower levels (Paap & Roske-Hofstrand, 1988). A major factor influencing success in finding information is the user’s understanding of the category names and partitions used by designers (Dumais & Landauer, 1983).

Content and structure unfamiliarity increases the amount of time necessary to find information (Somberg & Picardi, 1983). The disorientation that occurs when users become lost is so widespread and serious that Edwards and Hardman (1989) call it “hypertext’s major limiting factor” (p. 62). A cognitive explanation of reductions in performance is that users spend some of their limited processing capabilities on navigation that could otherwise be spent on attending to comprehension of the content (Thuring et al., 1995). Thus site understandability needs to be as high as possible so that the cognitive overhead does not make the site unusable.

Site understandability is more difficult to provide in very unfamiliar sites. Unfamiliar category names might require that users locate information through trial and error, which becomes especially difficult when menu labels are ambiguous or otherwise difficult to evaluate.
When dealing with unfamiliar categories it is possible for users to spend a great deal of time lost in the hierarchy of the system (Norman & Chin, 1988; Robertson et al., 1981). Paap and Roske-Hofstrand (1988) suggested over 10 years ago that user familiarity with category names should influence the design of hierarchical structures. When familiarity with categories is expected to be low, designers should use more general categories or switch to another technique such as alphabetical listing.

Ambiguity of the terms used by Web developers could make the sites more difficult to navigate. Quite often, firms organize Web sites according to their own internally-generated product line categories. For example, one very large consumer electronics firm once included three paths to its camcorders. The links were “camcorders,” “ultracams,” (name disguised) and “professional camcorders.” All were camcorders, but in different price, size, and/or feature groupings. Within each category, a customer would then need to select the desired format (such as 8mm, Hi8mm, Digital8, and DV). Even if a customer wanted to select a particular DV unit, he or she would have difficulty discerning the proper path in the site’s hierarchy to see the options. The extremely slow speed of each page of this firm’s site (8-10 seconds per page even with a LAN connection) exacerbated user frustration by imposing a very high waiting penalty for each mistake in the trial-and-error navigation.

The use of well-labeled, meaningful categories is one of the determinants of effective e-commerce Web sites (Turban & Gehrke, 2000) and of Web site ease of use (Lederer et al., 2000). We propose the following hypotheses regarding familiarity:

*H3a: A Web site with familiar structure and content will result in more favorable user attitudes about the site than will an unfamiliar Web site.*

*H3b: A Web site with familiar structure and content will result in higher user performance than will an unfamiliar Web site.*
**Interactions**

The model asserts that speed, depth, and familiarity have not only direct but also interactive effects on attitudes, performance, and finally, behavioral intentions. Though somewhat unusual, we also hypothesize all possible interactions.

**Depth and Familiarity**

Increasing the depth of a site will increase the number of clicks needed to reach the target node. When a site is unfamiliar, a larger number of clicks will be needed as well. The two factors are more than merely additive, as the formulas below will reveal.

A familiar site with $L$ levels to traverse (from the main index page to a bottom-level page) would require an average of $C$ clicks to reach the bottom node as follows:

$$C_{\text{fam}} = L$$

By definition, a familiar site causes no user errors in navigation; therefore the number of clicks to reach a bottom-level page is equal to the number of levels the user must traverse, with no backtracking. For example, if there are 5 levels from top to bottom, a user who starts at level 1 needs to traverse 4 levels to reach level 5, the goal.

An unfamiliar site, however, would at worst require a “brute force” systematic (next neighbor) search with backtracking. Such a search would require many more clicks to reach the correct bottom node. The expected number of clicks to reach the desired node is a function of the number of links per page and the number of levels:

$$C_{\text{unfam}} = \left( \sum_{i=1}^{L-1} n^i \right) + 1$$

where:

$n = \text{the number of links per page (assumed to be constant throughout)}$

$L = \text{the number of levels to traverse}$
This formula requires more explanation. Counting from the level under the home page, at each level $i$, there are $n^i$ intermediate pages. In a “brute force” search, the user would traverse each of these pages, hoping to find a direct link to the bottom-level node sought. Each unsuccessful search would normally stop at level $L-1$, that is, one level above the bottom level, in the absence of the desired link. In other words, if a user is looking for a certain model camcorder, and arrives at a page including only links to the pages of other models, the user would not need to load each page containing the incorrect model, so he or she moves on. Ultimately, when the user finds a link to the correct bottom page, he or she would need to click one last time to get there, hence the +1 at the end of the formula.

Traversing through the structure requires backtracking after each “dead end,” suggesting two clicks per intermediate page (one down and one back up). However, on average, a randomly-placed goal would be found halfway through the search. Therefore, the effect of backtracking is cancelled by finding the goal halfway through the task (multiplying by 2 clicks per page, and then dividing by 2), so the formula provides the average number of clicks to arrive at the goal.

If a site has 81 nodes (bottom level pages), two options might be to arrange them into 4 levels each with 3 links ($3^4=81$) or into 2 levels each with 9 links ($9^2=81$). The differences between the expected number of clicks in a familiar and unfamiliar site using each option are striking (see Table 1), indicating strong potential for interaction.

<table>
<thead>
<tr>
<th></th>
<th>Deep (4 levels to traverse, with 3 links each)</th>
<th>Broad (2 levels to traverse with 9 links each)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar Site</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Unfamiliar Site</td>
<td>40</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 1: Average Number of Clicks Predicted for Navigating to a Given Node
The pattern in Table 1 supports the following hypotheses:

\[ H4a: \text{There will be an interaction between familiarity and depth on user attitudes.} \]
\[ H4b: \text{There will be an interaction between familiarity and depth on user performance.} \]

**Depth and Speed**

Table 1 above illustrates that if a site has a deep structure, there will be a need to select many times the number of links (2 to 4 times in that example) to reach the desired bottom-level page than if a site has a broad structure. In a slow, deep site, each click brings yet another delay, presenting further opportunity for frustration and confusion. In short, Hypotheses 5a and 5b state that a “broad” design can partially make up for slow speed.

\[ H5a: \text{There will be an interaction between speed and depth on user attitudes.} \]
\[ H5b: \text{There will be an interaction between speed and depth on user performance.} \]

**Familiarity and Speed**

If a user has no idea where a bottom-level page is located, a “brute force” strategy (described earlier) might be needed to navigate to the proper page, leading to many times the number of clicks (5 to 10 times in that example) than when the path is either obvious or known. When response time is very poor, the “penalty” (delay) of moving from page to page becomes more dramatic. Indeed, Galletta et al. (2002) found that attitudes “bottomed out” sooner for unfamiliar sites than familiar sites. Attitudes ceased significant reductions after exceeding 2 seconds per page load for the familiar site and 6 seconds per page load for the unfamiliar site.

\[ H6a: \text{There will be an interaction between familiarity and speed on user attitudes.} \]
\[ H6b: \text{There will be an interaction between familiarity and speed on user performance.} \]

**Speed, Depth, and Familiarity**

Building on Table 1’s two-way interaction between depth and familiarity, the Web site’s loading speed adds another important dimension to the picture. As the delay with each click in-
creases, the differences among the cells increase dramatically. For example, if pages loaded instantaneously, the expected delay (for all clicks to the goal) would be zero in all cells (see Table 2), minimizing (but perhaps not eliminating completely) their differences. However, if a site incurred an 8-second delay, a brute-force search in the deep site would impose an expected delay of 320 seconds (5.3 minutes), while a brute-force search in the broad site would impose a delay of only 80 seconds (1.3 minutes). In a familiar site, the user would only need to click once at each level, with a delay of 32 seconds and 16 seconds in the deep and broad site, respectively.

\textbf{Table 2: Delay Predicted for Navigating to a Given Node, 0 or 8 second delay}

<table>
<thead>
<tr>
<th></th>
<th>Deep (4 levels to traverse, with 3 links each)</th>
<th>Broad (2 levels to traverse with 9 links each)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar Site</td>
<td>0 second delay / 8-second delay</td>
<td>0 second delay / 8-second delay</td>
</tr>
<tr>
<td>Unfamiliar Site</td>
<td>0 seconds / 32 seconds</td>
<td>0 seconds / 16 seconds</td>
</tr>
<tr>
<td></td>
<td>0 seconds / 320 seconds</td>
<td>0 seconds / 80 seconds</td>
</tr>
</tbody>
</table>

Therefore, the interaction increases in importance dramatically as delays increase as hypothesized below:

\textit{H7a: There will be an interaction between speed, familiarity, and depth on user attitudes.}
\textit{H7b: There will be an interaction between speed, familiarity, and depth on user performance.}

\textbf{Attitudes and Behavioral Intentions}

One of the key issues in designing web sites is getting potential users to visit and to return for future visits. Recent research suggests that the use of behavioral intentions is appropriate in a web context (Song & Zahedi, 2001). In this context a web user’s behavioral intentions (to return to the site) serves as an excellent summary variable that indicates design success. Fishbein and Ajzen (1975) theorize attitudes to be a central antecedent of intentions for future behavior. The relationship between attitudes and behavior has seen significant attention (Ajzen, 1991; Ajzen, 2001). MIS research in the individual acceptance of technology (Davis, 1989; Venkatesh
& Davis, 2000; Taylor & Todd, 1995) follows from this stream of research. Based on this work, most researchers would expect a direct and positive relationship between attitudes and behavioral intentions. Therefore we hypothesize that:

\[ H8a: \text{More favorable user attitudes will lead to more favorable behavioral intentions about the site.} \]

**Performance and Behavioral Intentions**

The HCI literature has focused substantial attention on performance as the central dependent variable for users of technology. System designs in both hardware and software arenas have been evaluated with user-system performance as the “ultimate concern” (Card et al., 1983, p. 404). Although their focus has been on choosing design factors based on their contributions to performance, HCI researchers have used quantitative performance models to make multi-million dollar adoption decisions (e.g., Gray et al., 1993).

Users will adopt strategies that maximize overall productivity (Teal & Rudnicky, 1992). Shneiderman (1998) states that “If delays are long, users will seek alternative strategies…perform other tasks…or plan ahead in their work” (pp. 360-361). Delays are not as problematic, however, when tasks involve time-consuming complex human problem-solving (Grossberg et al., 1976), where the delay is a much smaller portion of the overall task and would obviously not contribute much to productivity losses.

Over the last two decades, hundreds of studies in HCI have provided quantitative analysis of user-system performance for choosing (i.e., adopting) the best technological approach, in a variety of situations. Design of both software and hardware almost always involves such analysis, whether the designer is considering how formulas are to be entered into a spreadsheet, how an external mouse compares to a laptop pointing device, or how much time is saved by using a
pupil tracking system.

In the MIS literature, Compeau and colleagues have posited a link between performance and behavior for several years, based on Social Cognitive Theory (SCT) (see Bandura, 1977 and Compeau & Higgins, 1995, for a review of SCT). In a recent study by Compeau et al. (1999), the two strongest predictors of usage behavior were performance and attitudes (in both cases \( r = .25 \) and \( p < .001 \)). Therefore, there is significant evidence for suspecting a link between performance and behavioral intentions, and we present the final hypothesis:

\[ H8b: \text{Higher user performance will lead to more favorable behavioral intentions about the site.} \]

**Dependent Variables: Summary**

To summarize, attitudes and performance are theorized to be important antecedents of behavioral intentions, when a user has other alternatives to using the system or browsing on a particular site. In this study, we focus on speed, site depth, and familiarity as important Web site design factors, serving as direct and interacting antecedents of user attitudes and performance, and ultimately behavioral intentions.

**Research Methodology**

The study was conducted in an experimental setting to control speed, site depth, and familiarity and measurement of outcome variables. Two levels of each factor were provided to represent a relatively “high” and relatively “low” presence of the factor based on previous research, and we will use that shorthand to refer to each condition in spite of the somewhat arbitrary nature of each label, being consistent with previous research. Two levels of each of three factors provided a 2x2x2 design with two between-subjects factors (speed and depth) and one within-subjects factor (familiarity). We employed a completely counterbalanced, fully factorial design,
providing all 32 combinations of order, speed, depth, and familiarity.³

**Operationalization of Variables**

**Speed**

Speed, the central factor in this study, was manipulated by using Javascript to provide a constant 8-second delay per page for the slow site, with no such delay for the fast site. As discussed earlier, a “long” delay is a subjective judgment. An eight second delay is about the middle of the range of what others have considered serious (Hoxmeier & DiCesare, 2000; Ramsay *et al.*, 1998; Zona, 1999; Shneiderman, 1998; Kuhmann, 1989), and has been shown to be well past the patience levels of laboratory subjects (Galletta *et al.*, 2002). The eight second delay also worked well in our laboratory situation because a longer delay might not be justified by the simplicity of the pages (Sears & Jacko, 2000).

A manipulation check for the speed variable asked subjects to evaluate the speed of displaying the Web pages. On a 7-point Likert scale, “fast” subjects, on average, rated the speed as 6.0 (standard deviation 1.2) and “slow” subjects rated the speed as 2.2 (standard deviation 1.3). The difference was significant (t=312.3; one-tailed p<.001).

**Familiarity**

Two artificial Web sites were created. “Pete’s General Store” and “A.C.T. Systems” served as the “familiar” and “unfamiliar” sites, respectively. The “familiar” site contained products found in grocery and/or hardware stores almost anywhere, to ensure that subjects would recognize these products and categories of products (for example, a top level category was “Health Care Products”). In contrast, the “unfamiliar” site contained fictitious software products

---

³ To prevent any learning effects in the second task, we counterbalanced completely the order for within-subjects factors. 32 combinations = 2 familiarity orders (familiar-unfamiliar and unfamiliar-familiar) x 4 speed conditions (fast-fast, fast-slow, slow-fast, and slow-slow) x 4 depth conditions (deep-deep, deep-broad, broad-deep, broad-broad).
and accessories and their categories were meaningless (for example, a top level category was “Novo Products”). Subjects could not make use of previous experience in searching through the unfamiliar site. Both sites were complete with brief product descriptions, prices and images.

A manipulation check for the familiarity variable asked subjects to evaluate their level of familiarity with the subject matter in the Web site. On a 7-point Likert scale, subjects on average rated the familiar site 5.4 (standard deviation 1.4) and unfamiliar site as 2.2 (standard deviation 1.4) on a 7-point Likert scale. The difference was significant (t=21.3; one-tailed p<.001).

**Depth**

The experimental Web sites were created with two different hierarchical depths for the 81 bottom level pages. A “broad” structure was created with 2 levels to traverse (9 hyperlinks per page), and a “deep” structure contained 4 levels to traverse (3 hyperlinks per page). These levels were chosen after considering previous studies (Norman & Chin, 1988; Landauer & Nachbar, 1985; Kiger, 1984; Miller, 1981), which used very similar schemes.

A manipulation check for the depth variable asked subjects to evaluate the number of further choices or links on each page visited in the Web site. On a 7-point Likert scale, subjects on average rated the broad site 4.1 (standard deviation 1.9) and the deep site 3.9 (standard deviation 1.6). Although in the expected direction, the difference was not significant (t=1.0; one-tailed p<.16), suggesting that either the condition or the manipulation check measure was inadequate.

Further testing on the manipulation check for depth revealed that users had trouble judging “absolute” depth or breadth in this setting. Because the effect of depth has greatest implications in the unfamiliar site, effectively multiplying the number of required clicks by at least 5, we omitted the scores of all subjects who went from the unfamiliar to the familiar sites, assuming that the extra two clicks in the deep familiar site would be too trivial to comprise a meaningful
difference in perceived depth. For the subjects who browsed the unfamiliar site as their second site, the mean score for the broad site was 4.6 (standard deviation 1.8) and for the deep site was 3.9 (standard deviation 1.9), differing significantly on the depth variable \( t=1.7; \) one-tailed \( p<.046 \). Thus, there is some evidence that the subjects could perceive and judge the depth/breadth manipulation, though not as strongly as the other factors.

**Dependent Variables**

Table 3 provides a summary of the number of items, types of items, and reliability analysis for each scale.

**Table 3: Instrument Reliabilities**

<table>
<thead>
<tr>
<th>Dependent Measure</th>
<th>Number of Items</th>
<th>Type of Items</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes</td>
<td>6</td>
<td>9-point scale</td>
<td>alpha = .95</td>
</tr>
<tr>
<td>Performance</td>
<td>9</td>
<td>Dichotomous</td>
<td>KR-20 = .88</td>
</tr>
<tr>
<td>Behavioral Intentions</td>
<td>2</td>
<td>7-point scale</td>
<td>alpha = .94</td>
</tr>
</tbody>
</table>

**Attitudes**

Attitudes about the sites were measured averaging the responses to a set of seven 9-point Likert-type questions adapted from Part 3 of the long form of the QUIS (Questionnaire for User Interaction Satisfaction) (Shneiderman, 1998, p. 136), which has demonstrated reliability and validity (Chin et al., 1988). We dropped one item to maximize Cronbach’s alpha, which was .95.

**Performance**

Performance was operationalized using an average score for nine search tasks that were developed for this study. Whether the site was familiar or unfamiliar, for accomplishing their search tasks the questions were identical except for the search object names, which were general store products in one site and software products in the other site. The structure was identical in each site; subjects needed to visit exactly the same position of the structure in each site.
only differences between the familiar and unfamiliar sites were the products themselves.

The nine tasks required browsing the site for facts (such as prices or packaging) about various products. Subjects were asked to fill in the answer after reaching the proper Web page, and received 1 point for each correct answer. Short, dichotomous instruments are not normally subjected to reliability analysis, so the results of such analysis should not be viewed with the same expectations as Likert-type instruments or dichotomous instruments with over 20 items (Nunnally, 1978). Nevertheless, the statistic of the Kuder-Richardson-20, or KR-20 test (analogous to alpha, but for dichotomous items), was quite high in this study (.88).

**Behavioral Intentions**

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

**Subjects**

Undergraduate business majors enrolled in sections of an upper-division undergraduate management information systems course at a large Northeastern U.S. university were invited to participate in the study. Students were considered to be appropriate for this study because the materials were designed to tap “invariant” (Simon, 1990) aspects of Web use. In addition, Voich (1995) found students to be particularly representative of values and beliefs of individuals employed in a variety of occupations.

Of 191 students invited, 177 (93%) volunteered to participate. The first 160 served to fill 5 completely counterbalanced groups of 32 each (77 females, 83 males), and the last 17 subjects
were discarded. No cash incentives were provided, but subjects were given the opportunity to participate during one of their class periods. The within- and between-subjects design provided 40 subjects for each of the 8 cells.

A pre-test employing a separate group of 32 subjects from the same sampling pool (1 for each of the 32 combinations of conditions) in a previous term helped us refine the instruments and procedures. Those data are not included in our results.

Subjects were randomly assigned to one of the 32 counterbalanced conditions, and therefore, randomly assigned to treatments. Both multivariate and univariate ANOVAs were used as tests of randomization using three selected demographics: gender, computer efficacy, and computer experience. Computer efficacy was measured with the 10-item instrument by Compeau and Higgins (1995) and computer experience was measured with a 15-item instrument adapted from the technical skills and IS product knowledge sections of Nelson’s (1991) instrument. Both adaptations followed that of Hartzel (1999), and our results show alpha of .90 for computer efficacy and .91 for computer experience.

In general, ANOVA revealed no differences among the 8 cells on gender, computer experience, and computer efficacy. Of the 21 2x2x2 tests (3 main effects and 4 interactions for each of the three demographic variables), only one difference was significant: computer efficacy differed along a 2-way interaction of speed and depth (F=4.7; p=.032; 1,149 df). Although one test out of 21 is expected to be significant at the .05 level based purely by chance, it would be of greater concern if it were highly significant and on a main effect.

Procedure

The Web sites were created and their 32 combinations were written on CDs to precisely
control the browser’s response time. A large computer laboratory containing 46 identical Pentium II computers and 17” XGA screens further controlled the subjects’ environments. When the subjects arrived, they were asked to sign an “informed consent” form, and were told that participation was voluntary and that they could leave the study at any time. They were then instructed to enter the randomly-assigned code number (1-32) printed on their packet to trigger the system to load the correct treatment condition. Automatic logs (provided by software running hidden in the background) and manipulation check questions described earlier provided assurance that the correct number was entered.

A very small sample site provided a “warm-up.” Then the subjects moved on to complete the required 9 tasks for each of the two main sites. At the conclusion of each main site, the subjects were instructed to close the browser window and complete the questions addressing attitudes and behavioral intentions. Although individual times varied, on average subjects completed the experiment in about an hour.

Results

Because the correlation between performance and attitudes was .51 (p<.001), our analysis of each hypothesis from H1 through H7 began with a multivariate approach followed by a univariate ANOVA for each dependent variable; each MANOVA test was significant.

Tables 4 and 5 provide the means for each dependent variable and Table 6 provides the multivariate tests of H1 through H7. The tests reveal that every main and every possible interaction effect is significant (p<.005). Therefore, univariate testing was performed for each hypothesis.

4 The CD, as well as all instruments and materials, are available from the first author on request.
Table 4: Cell Means for Attitudes

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Slow Mean</th>
<th>Standard Deviation</th>
<th>Fast Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>1.8</td>
<td>1.1</td>
<td>2.8</td>
<td>1.6</td>
</tr>
<tr>
<td>Familiar</td>
<td>5.0</td>
<td>2.0</td>
<td>6.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Broad</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>2.5</td>
<td>1.5</td>
<td>4.2</td>
<td>2.2</td>
</tr>
<tr>
<td>Familiar</td>
<td>6.5</td>
<td>1.7</td>
<td>7.0</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Table 5: Cell Means for Performance

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Slow Mean</th>
<th>Standard Deviation</th>
<th>Fast Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>.42</td>
<td>.29</td>
<td>.80</td>
<td>.30</td>
</tr>
<tr>
<td>Familiar</td>
<td>.95</td>
<td>.09</td>
<td>.98</td>
<td>.05</td>
</tr>
<tr>
<td>Broad</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>.81</td>
<td>.23</td>
<td>.93</td>
<td>.17</td>
</tr>
<tr>
<td>Familiar</td>
<td>1.0</td>
<td>0</td>
<td>1.0</td>
<td>.02</td>
</tr>
</tbody>
</table>

Table 6: Overall MANOVA: Both Dependent Variables (for H1-H7)

<table>
<thead>
<tr>
<th>Factor</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>36.4</td>
<td>.000</td>
</tr>
<tr>
<td>Familiarity</td>
<td>221.0</td>
<td>.000</td>
</tr>
<tr>
<td>Depth</td>
<td>36.6</td>
<td>.000</td>
</tr>
<tr>
<td>Speed * Familiarity</td>
<td>17.2</td>
<td>.000</td>
</tr>
<tr>
<td>Speed * Depth</td>
<td>6.3</td>
<td>.002</td>
</tr>
<tr>
<td>Depth * Familiarity</td>
<td>15.7</td>
<td>.000</td>
</tr>
<tr>
<td>Speed * Familiarity * Depth</td>
<td>8.3</td>
<td>.000</td>
</tr>
</tbody>
</table>

Individual ANOVAs tests were run for both variables, and the amount of variance explained by each analysis (including interaction effects) is .569 for attitudes and .498 for performance (adjusted $R^2$).

Means for the dependent variables, and results of hypothesis tests for the main effects are presented in Tables 7-9. All 6 tests are significant (in the predicted direction) at the .001 level,
providing strong support for the main effects predicted by the model.

**Table 7: The Speed Main Effect**

<table>
<thead>
<tr>
<th>Dependent Measure</th>
<th>Slow</th>
<th>Fast</th>
<th>Difference (F, p) at 1,312 df</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Mean</td>
</tr>
<tr>
<td>Attitudes</td>
<td>4.0</td>
<td>2.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Performance</td>
<td>.80</td>
<td>.30</td>
<td>.92</td>
</tr>
</tbody>
</table>

**Table 8: The Depth Main Effect**

<table>
<thead>
<tr>
<th>Dependent Measure</th>
<th>Deep</th>
<th>Broad</th>
<th>Difference (F, p) at 1,312 df</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Mean</td>
</tr>
<tr>
<td>Attitudes</td>
<td>4.0</td>
<td>2.5</td>
<td>5.1</td>
</tr>
<tr>
<td>Performance</td>
<td>.79</td>
<td>.31</td>
<td>.93</td>
</tr>
</tbody>
</table>

**Table 9: The Familiarity Main Effect**

<table>
<thead>
<tr>
<th>Dependent Measure</th>
<th>Unfamiliar</th>
<th>Familiar</th>
<th>Difference (F, p) at 1,312 df</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Mean</td>
</tr>
<tr>
<td>Attitudes</td>
<td>2.8</td>
<td>1.8</td>
<td>6.3</td>
</tr>
<tr>
<td>Performance</td>
<td>.74</td>
<td>.31</td>
<td>.98</td>
</tr>
</tbody>
</table>

**Main Effects**

Attitudes and performance were significantly more positive in the faster site, broader site, and familiar site, providing clear support for Hypotheses H1a, H1b, H2a, H2b, H3a, and H3b.

**Interaction Effects**

**Depth and Familiarity**

Hypotheses H4a and H4b predict that the positive effects of breadth are more important for an unfamiliar site than for a familiar site (and conversely, the negative effects of depth are
more severe for an unfamiliar site than a familiar site). Although the interaction effect was not significant for attitudes (H4a), the effect seems to hold very clearly, and precisely in the expected direction, for performance (H4b; F=31.4; p<.001).

**Depth and Speed**

Hypotheses H5a and H5b predict that the positive effects of breadth (rather than depth) are enhanced more for a slower site than for a faster site. As in the analysis of the previous interaction, such an interaction was found for performance (H5b; F=12.57; p<.001) but not attitudes (H5a). The means are in the expected direction, demonstrating that breadth indeed reduces the performance disadvantage of a slow site.

**Familiarity and Speed**

Hypotheses H6a and H6b predict that the positive effects of familiarity are enhanced more for a slower site than for a faster site. Once again, the interaction effect on performance was significant (H6b; F=34.5; p<.001) but not on attitudes (H6a). The results demonstrate a strong interaction, as predicted, indicating that familiarity reduces the performance disadvantage of a slow site.

**Speed, Depth, and Familiarity**

Hypotheses H7a and H7b predict that the combined positive effects of familiarity and breadth are enhanced more for a slower site than for a faster site. The results support both H7a and H7b, demonstrating a three-way interaction for both attitudes (F=17.9; p<.011) and task performance (F=7.98; p<.005). The cell means reveal that familiarity and breadth interact to reduce the performance and attitudinal disadvantage of a slow site.

**Antecedents of Behavioral Intentions**

H8a and H8b state that behavioral intentions can be predicted by attitudes and perform-
Intercorrelations among the constructs indicated general support for both hypotheses (p < .001), with a .751 correlation between attitudes and behavioral intentions and a .392 correlation between performance and behavioral intentions.

To determine the relative contribution of each variable in predicting behavioral intentions, a multiple regression approach was also used. The regression model included intentions as a dependent variable and attitudes and performance as independent variables, and was found to explain significant variance in behavioral intentions (adjusted R² = .561). However, only the attitude measure contributed significantly to the model (t = 17.3; p < .001). In a stepwise regression, only attitudes entered the model (adjusted R² = .562).

Table 13 provides a summary of the Hypotheses and results. The results indicate that, perhaps unsurprisingly, there are indeed powerful main effects from Web site speed, depth, and familiarity on user attitudes and performance in searching through the site. However, beyond those intuitive main effects, for performance we found strong two-way interactions between depth and familiarity, depth and speed, and between familiarity and speed as expected. A strong three-way interaction among all three factors was found for both performance and attitudes. Finally, attitudes appear to be a powerful antecedent of behavioral intentions.

Interestingly, the effects on performance were consistently stronger in every main and interaction test than the effects on attitudes. Meanwhile, 2-way interaction effects were completely absent for attitudes, and while most of the hypotheses received strong support from manipulating the three factors, the strongest findings are in the arena of performance. Simultaneously, attitudes are the strongest predictor of behavioral intentions, completely overwhelming any additional impacts of performance. Therefore, a focus only on performance or only on attitudes appears to be rather incomplete.
Table 13: Summary of Findings

<table>
<thead>
<tr>
<th></th>
<th>Expectation</th>
<th>Effect of Factors</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a</td>
<td>Attitudes: Fast &gt; Slow</td>
<td>Main</td>
<td>Supported</td>
</tr>
<tr>
<td>H1b</td>
<td>Performance: Fast &gt; Slow</td>
<td>Main</td>
<td>Supported</td>
</tr>
<tr>
<td>H2a</td>
<td>Attitudes: Broad &gt; Deep</td>
<td>Main</td>
<td>Supported</td>
</tr>
<tr>
<td>H2b</td>
<td>Performance: Broad &gt; Deep</td>
<td>Main</td>
<td>Supported</td>
</tr>
<tr>
<td>H3a</td>
<td>Attitudes: Familiar &gt; Unfamiliar</td>
<td>Main</td>
<td>Supported</td>
</tr>
<tr>
<td>H3b</td>
<td>Performance: Familiar &gt; Unfamiliar</td>
<td>Main</td>
<td>Supported</td>
</tr>
<tr>
<td>H4a</td>
<td>Attitudes: Interaction between Depth &amp; Familiarity</td>
<td>2-way</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H4b</td>
<td>Performance: Interaction between Depth &amp; Familiarity</td>
<td>2-way</td>
<td>Supported</td>
</tr>
<tr>
<td>H5a</td>
<td>Attitudes: Interaction between Depth &amp; Speed</td>
<td>2-way</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H5b</td>
<td>Performance Interaction between Depth &amp; Speed</td>
<td>2-way</td>
<td>Supported</td>
</tr>
<tr>
<td>H6a</td>
<td>Attitudes: Interaction between Familiarity &amp; Speed</td>
<td>2-way</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H6b</td>
<td>Performance: Interaction between Familiarity &amp; Speed</td>
<td>2-way</td>
<td>Supported</td>
</tr>
<tr>
<td>H7a</td>
<td>Attitudes: Interaction between Speed, Depth, &amp; Familiarity</td>
<td>3-way</td>
<td>Supported</td>
</tr>
<tr>
<td>H7b</td>
<td>Performance: Interaction between Speed, Depth, &amp; Familiarity</td>
<td>3-way</td>
<td>Supported</td>
</tr>
<tr>
<td>H8a</td>
<td>Attitudes as an antecedent of Behavioral Intentions</td>
<td>n/a</td>
<td>Supported</td>
</tr>
<tr>
<td>H8b</td>
<td>Performance as an antecedent of Behavioral Intentions</td>
<td>n/a</td>
<td>Not Supported</td>
</tr>
</tbody>
</table>

Summary and Conclusions

The slow speed of World-Wide Web sites is a commonly-discussed problem that has several potential causes and effects. Because the causes could occur at many locations such as the server (the site itself), network congestion, global demand for visiting the site, and the user’s local environment, our current problems are likely to continue for several more years because no single solution can provide an effective remedy.

The problem of speed is magnified when users need to load several pages or backtrack as they browse a site for a particular target. This need is quite often a result of a site that has unfamiliar structure or content, leading to errors, or one that is deep, with many intermediate pages.

The literature suggests that each of the three factors, speed, familiarity, and depth, has strong user performance or preference outcomes: slow speed, unfamiliar structure and content, and deep
sites (with many levels) can have negative effects compared to fast, familiar, and broad sites.

Our model predicted all possible main and interactive effects of the three factors on users’ attitudes and performance on nine search tasks. While similar main effects have been found in many contexts in studies performed over the last 30 years, there has been little research to discover to what extent the factors interact. The predicted interactions form the heart of this study, and state that difficulties of slow loading speed can be magnified by high unfamiliarity and site depth beyond the linear impacts of each of the three factors by themselves. Also important are the predicted two-way interactions, based on the wide variations in the expected number of steps (clicks) needed to browse broad and deep, familiar and unfamiliar sites for target information. Formulas are provided to make these predictions. The model also predicted that behavioral intentions would, in turn, be influenced by both attitudes and performance.

MANOVA and ANOVA were performed on data from 160 subjects, randomly assigned to treatments with high (8 second) or low delay, familiar or unfamiliar links and content, and broad (2 levels/9 links per page) or deep (4 levels/3 links per page) structure. The 3-way ANOVAs explained 57% of the variance in attitudes and 50% of the variance in performance.

As expected, all main effects were highly significant on both of the hypothesized outcome variables. Attitudes and performance on the search task were strongly degraded by slow speed, unfamiliar structure and content, and deep structure. Those results agree with decades of reasonably consistent and unequivocal work in the HCI literature, and provide some assurance that the manipulations chosen in this study did have strong effects on the outcomes we examined.

The interaction effects seemed to have the strongest effects on user performance. All possible 2-way interactions among the factors showed significant effects on performance, as well as the 3-way interaction. Only the 3-way interaction was significant for attitudes. Although it is
relatively rare to theorize and find 3-way interactions, it appears that the three factors indeed have a synergistic relationship, even to the extent that they seem to overshadow some of the 2-way interactions for attitudes.

Although behavioral intentions correlated significantly with both attitudes and performance, a multiple regression revealed that the effects of attitudes appeared to overshadow completely the effects of performance on behavioral intentions.

This study is meant to provide implications for both researchers and practitioners. The research model offers a contingency framework that goes beyond the main effects described in previous related research, and a richer basis for conceptualizing and calculating strong performance impacts of all possible interactions among the factors. We suggest that it is inadequate to focus only on one of the three factors in isolation, and future research needs to consider multiple factors for more realistic predictions and more complete analysis of the phenomenon of slow Web sites. While our model appears to explain a large amount of variation in attitudes, task performance, navigation performance, and time performance, there is ample opportunity to focus on other outcomes and other antecedents which might lead to greater understanding of the usability of Web sites. Also, while we were able to demonstrate correlations between behavioral intentions and both attitudes and task performance, other studies should focus very closely on these three constructs to better understand their relationships under different situations.

Practitioners should be advised to pay close attention to each of the three individual factors within their control to the greatest extent possible. Page loading time should be minimized, links to bottom level pages should use the most familiar possible terminology, and structural pages (leading to the bottom level pages) should have more, rather than fewer, links per page.

More importantly, however, although practitioners are not usually made aware of interac-
tions among experimental factors, the ones in this study provide some clear advice:

- When it is difficult or impossible to expect a reasonable amount of familiarity in target users, then a great deal of effort should be expended to maximize breadth and speed (either by minimizing graphic content or assuring proper server configuration, redundancy, and/or bandwidth).
- When familiarity is expected to be high, pages can be made to be more attractive by using graphics, and the site can be made with greater depth than might otherwise have been considered advantageous.
- When it is likely that speed will be an issue because of the need for graphical content or likely bottlenecks, then breadth should be maximized and the site’s structure should employ very familiar, clear terminology.

A final concern for practitioners is that this study shows that the widely-touted success measures of the mid-1990s need to be refined. The number of page visits and time spent on a site are only part of the picture; while a confusing and deep structure can provide larger numbers of page views, intentions to revisit are likely to suffer due to degradations of attitudes.

While analysis of single factors such as those investigated in this study are important for establishing foundations for understanding basic design phenomena, analysis of interactions provide the basis for simultaneously more thorough understanding for researchers and more practical heuristics for designers. Refining and extending such interactive models would be a promising avenue for future work.

References


Connolly, P.J., “Growing Pains Won’t Slow Down the Broadband Revolution,” *Infoworld*, November 6, 2000, 66 and 68.


In-Stat/MDR, 2002; http://www.instat.com/


Nelson, R.R. “Educational Needs as Perceived by IS and end-user personnel: A Survey of

Nielsen, J. “Who Commits The "Top Ten Mistakes" of Web Design?,”

Nielsen, J. “The Top Ten New Mistakes of Web Design,”


Shneiderman, B. *Designing the User Interface*, Addison-Wesley, Boston, 1998.


Wonnacott, L. “When User Experience is at Risk, Tier 1 Providers can Help,” *Infoworld*, November 6, 2000, p. 76.
Zaphiris, P. and Mtei, L. “Depth vs Breadth in the Arrangement of Web Links,”


Zona Research Study, “The Need for Speed,”