A study on tolerable waiting time: how long are Web users willing to wait?

Citation: Nah, F. (2004), A study on tolerable waiting time: how long are Web users willing to wait? *Behaviour & Information Technology,* forthcoming

Fiona Fui-Hoon Nah
College of Business Administration
University of Nebraska-Lincoln
Lincoln, Nebraska 68588-0491
USA
Tel: (402) 472-6060
Fax: (402) 472-5855
Email: fnah@unl.edu

Abstract

Web users often face a long waiting time for downloading Web pages. Although various technologies and techniques have been implemented to alleviate the situation and to comfort the impatient users, little research has been done to assess what constitutes an acceptable and tolerable waiting time for Web users. This research reviews the literature on computer response time and users' waiting time for download of Web pages, and assesses Web users' tolerable waiting time in information retrieval. It addresses the following questions through an experimental study: What is the effect of feedback on users' tolerable waiting time? How long are users willing to wait for a Web page to be downloaded before abandoning it? The results from this study suggest that the presence of feedback prolongs Web users' tolerable waiting time and the tolerable waiting time for information retrieval is approximately 2 seconds.

1. Introduction

The World Wide Web (WWW) has become an important channel for information retrieval, electronic commerce and entertainment. However, long Web page download times have remained a major cause of frustration among Web users (Selvidge 1999, 2003). According to the findings of the surveys conducted by Lightner, Bose and Salvendy (1996) and the GVU (Graphic, Visualization and Usability) Centre at Georgia Institute of Technology (GVU, 1998), long download times have always been a major problem experienced by Web users. The survey by Pitkow and Kehoe (1996) also indicates that the most widely cited problem with using the WWW was that it took too long to download Web pages (i.e. 69% of respondents cited this problem). This problem is worsened by the exponential increase in the number of Web users over the years and the popularity of multimedia (e.g. video, voice) technology. This problem is so noticeable that Web users often equate the "WWW" acronym with "World Wide Wait"!

The WWW has become an important and popular information search tool. It provides convenient access to almost all kinds of information – from education to entertainment. It also makes global information available at our fingertips. Although the WWW is now accessible from mobile devices, usage and adoption rates are low due to the long download time and limited bandwidth available in the wireless environment. As noted earlier, the long waiting time for downloading Web pages is often not tolerable even in the wired environment. Due to the increasing and excessive use of multimedia data (i.e. audio and video clips) on Web pages, this concern is continuously growing. This problem of 'long download time' is relevant not only to Web users but also to the authors and designers of websites, as websites that take a long time to download are rarely or less frequently visited (Reaux and Carroll 1997). Hence, it is important for us to gain a more in-depth understanding of Web users' waiting behaviour. More specifically, the main questions of interest are: *How long are users willing to wait for downloading a Web page before abandoning it*? We will refer to this duration as the tolerable waiting time (TWT). *Does providing feedback during the wait prolong Web users' tolerable waiting time*?

The answers to the above questions are important for making decisions on hardware and software investments as well as Web page design and contents in order to provide acceptable download time to users. Network providers and website designers make such decisions based on their understanding of what constitutes an acceptable download time. Since the types of task (e.g. information retrieval, browsing, purchasing, downloading of files) may have an impact on users' level of tolerance, we will examine information retrieval task in this research, which is one of the most common tasks on the WWW.

The rest of this paper is organized as follows: The next section reviews the literature on users' TWT and formulates the hypothesis concerning the effect of feedback on Web users' TWT. Section 3 describes an experimental study to answer the questions in this research. The results of the study are reported in Section 4. Section 5 concludes with implications of the research and discussions for future research.

2. Literature on waiting time

According to Nielsen (1999: 67), download speeds are the 'single-most important design criterion on the Web'. Web users are constantly begging for faster page downloads (Nielsen 2000). Although long download time of Web pages has been a consistent problem encountered by Web users (Lightner *et al*. 1996, Pitkow and Kehoe 1996, Selvidge 1999, 2003), it is still controversial as to what constitutes an *acceptable* waiting time for a typical Web page download (Bailey 2001). Nielsen (1997) advocates the 10-second limit, while Zona Research (1999) recommends the 8-second rule. Selvidge's (1999) study shows that there is no difference in users' frustration levels between 1-second and 20-second delay, but a difference (with 1-second delay) was observed at 30-second delay. Other researchers propose the 2-second rule (Shneiderman 1984) and the 12-second rule (Hoxmeier and DiCesare 2000). The conflicting evidence in the literature was also highlighted and examined by Galletta, Henry, McCoy and Polak (2002), who observed decreases in performance and behavioural intentions at 4 seconds.

Interestingly and ironically, the average American users that use dial-up connections wait about 30 seconds the first time they look at a new Web page (Chen 2002)!

Web page download time is affected by the performance of the browser, the speed of the Internet connection, the local network traffic, the load on the remote host, and the structure and format of the Web page requested. In this research, we are not addressing the issue of how these different variables can be balanced or traded-off to produce an acceptable download time but rather, we are interested in finding out what constitutes an acceptable or, more exactly, tolerable download time for a typical Web user. Although acceptable or tolerable waiting time for Web page download can be defined from various perspectives (e.g. change in attitudes such as satisfaction or frustration, behavioural intentions such as intention to visit or not visit the site again, perceptions such as perceived waiting time, performance such as quality or accuracy, or user behaviour such as the act of abandoning a Web page), in this research, tolerable waiting time (TWT for short) is defined as *the amount of time users are willing to wait before giving up on the download of a Web page*.

2.1 Tolerable computer response time

This section reviews earlier research on computer response times where it is suggested that (Nielsen 1993):

(1) **0.1 second** is about the limit for having the user feel that the system is reacting instantaneously, meaning that no special feedback is necessary except to display the result.

(2) **1.0 second** is about the limit for the user's flow of thought to stay uninterrupted, even though the user will notice the delay. Normally, no special feedback is necessary during delays of more than 0.1 but less than 1.0 second, but the user does lose the feeling of operating directly on the data.

(3) **10 seconds** is about the limit for keeping the user's attention focused on the dialogue. For longer delays, users will want to perform other tasks while waiting for the computer to finish, so they should be given feedback indicating when the computer expects to be done.

Table 1 provides a summary of the literature on users' tolerance in waiting for computer response. Miller (1968) argued for the 2-second rule based on the theory of limitations in human short-term memory. According to Miller, short-term memory plays a critical role in human information processing; interference with short-term memory can occur when an individual senses an awareness of waiting after approximately 2 seconds. Thus, to stay uninterrupted in information processing, the 2-second guideline is recommended. For tasks where uninterrupted focus is critical, Nielsen (1995) suggests that computer response should be kept within one second. For other types of tasks, the threshold can go up to 10 seconds (Nielsen 1993). This is based on Miller's (1968: 268) proposition that 'a system with response delays of a standard ten seconds will not permit the kind of thinking continuity essential to sustained problem solving'.

Shneiderman (1984) reviewed the literature on computer response time and recommends that the computer should respond to users within two seconds. Shneiderman cited Youman's study where it was found that users' reactions were changed from predominantly acceptable to predominantly unacceptable around 2 seconds. Shneiderman also suggests that the 2-second limit is appropriate for most online tasks.

[Insert table 1 about here]

Although the impact of system response time has been investigated in the context of computer terminal and personal computer use, only a few studies have examined it in the context of the Internet and the WWW. The next section will review the literature on users' TWT for Web page download.

2.2 Tolerable waiting time for Web page download

The TWT for downloading a Web page may depend on various factors such as level of experience and age of users, individual user's characteristics (i.e. propensity to wait), task type, expected content of the Web page, expected download time, and information available about the wait. Such variability and its associated research challenge should not deter us from studying the ‘waiting time’ phenomenon and Web users’ waiting behaviour. This section reviews the literature on Web users’ TWT and presents a summary of the findings in table 2.

[Insert table 2 about here]

Ramsay, Barbesi and Preece (1998) found that page loading delays (from 2 seconds to 2 minutes) had strong effects on users’ perceptions of websites. The results indicate that Web pages that were downloaded faster were perceived to be more interesting than the slower ones. Their results also suggest that 41 seconds is the cut-off for long delays. On the other hand, Selvidge’s (1999) study suggests a threshold of 30 seconds. In the study, the effects of 1-, 30- and 60-second delays produced significant differences in both performance and frustration levels whereas no statistical difference was observed between the effects of 1-, 10- and 20-second delays.

According to Nielsen (1995, 1996), Web users may be willing to tolerate up to 15 seconds for a Web page download. Even though traditional human factor guidelines suggest that 10 seconds is the maximum response time before computer users lose interest (Miller 1968, Nielsen 1993, 1997), Nielsen (1995, 1996) suggests that 15 seconds is considered tolerable as Web users have been ‘trained to endure so much suffering that it may be acceptable to increase the limit value to 15 seconds’.

Hoxmeier and DiCesare (2000) also examined the relationship between system response time (0, 3, 6, 9, 12 seconds) and several users’ perception measures in browser-based applications. Their results show that satisfaction decreases with increases in response time and the level of intolerance occurs at the 12-second response range. Galletta, Henry, McCoy and Polak (2002) examined delay times of 0, 2, 4, 6, 8, 10 and 12 seconds using an experiment. Their findings suggest that, ‘decreases in performance and behavioural 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'. Thus, the users' TWT is around 4 seconds (since waiting behaviour is more closely related to performance and behavioural intentions than attitudes).

The literature review indicates that although several studies have investigated the relationships between page loading latency and Web users' perceptions and attitudes, empirical research that investigates *actual* waiting behaviour of Web users is scarce (Selvidge 2003). The review also indicates that a wide range (from 1 second to 41 seconds) of TWT has been proposed. Although other research works have been done to study Web page loading latency, we only included those that provide specific recommendations concerning Web users' TWT.

In summary, it is unclear from the literature what constitutes a reasonable and acceptable waiting time for download of Web pages. Is it 1, 2, 5, 10 or 15 seconds? Unfortunately, there is no clear empirical evidence that supports any of these "magic numbers". The question will need to be answered through empirical investigations, such as the study described in this paper.

2.3 Effect of feedback on TWT

One of the fundamental principles of Nielsen's 10 usability heuristics (available at: http://www.useit.com/papers/heuristic/heuristic_list.html) is to provide visibility of system status. According to this principle, 'the system should always keep users informed about what is going on, through appropriate feedback within reasonable time'. As the wait or latency increases, negative emotional feelings, such as user anxiety, increase (Guynes 1988). The negative effects of waiting can be neutralized by effectively managing waiting experiences (Katz *et al*. 1991, Taylor 1994, Hui and Tse 1996, Dellaert and Kahn 1999). For example, feedback can be provided in the form of a moving status bar or waiting duration information.

Although both types of feedback are worth studying, this research will focus on the first type of feedback – moving status bar – and examine its effect on Web users' TWT. Two theoretical perspectives are used to explain the relationship between this type of feedback and Web users' TWT: the resource-allocation perspective and the uncertainty reduction perspective.

The resource-allocation perspective is presented by Zakay and Hornik (1991), who argue that feedback can distract a user's attention from the passage of time. According to their model, each time unit in a wait is cognitively recorded when a user pays attention to the passage of time. By providing feedback to the user, the user's mental activity is increased and thus, less attention is paid to the wait itself. Such filled time appears to pass more quickly than empty (unfilled) time (Gilliland *et al*. 1946, Katz *et al*. 1991, Taylor 1994), thus extending the TWT of Web users.

The uncertainty reduction perspective suggests that feedback reduces users' uncertainty concerning the wait. According to Taylor (1994), filling time can reduce uncertainty felt by the user by reducing boredom, tension, and its resulting anxiety. Stress increases when one is uncertain about the wait (Osuna 1985, Hui and Tse 1996). Providing feedback during the wait lowers the level of stress experienced by the Web users and reduces the uncertainty of the wait, which in turn help to prolong users' TWT (Hui and Zhou 1996, Weinberg 2000).

Both the resource-allocation and uncertainty reduction perspectives suggest that providing feedback during Web page download will extend Web users' TWT. Hence, the following hypothesis is proposed:

H1: Web users' TWT will be extended by providing feedback during Web page download.

The next section describes an experimental study that was carried out to assess the effect of feedback on Web users' TWT (i.e. the point at which a wait is surrendered or abandoned). This study differs from past empirical studies in that it focuses on understanding Web users' *actual* waiting behaviour instead of Web users' *perceptions* about page loading latency.

3. Research model, methodology and task

Although users' tolerance of a Web page download may vary for different types of tasks (such as information retrieval, online purchasing, downloading of files), in this study, users' TWT was studied in the context of purposeful browsing (i.e. focused search) as opposed to open-browsing. Given that users' TWT may be moderated by a number of factors, we controlled for task type (information retrieval), user characteristics (sophomore business major students who were savvy Web users), browser interface (specifically designed for the study), domain of information retrieval (information on software and hardware tools), and the specific Web pages that were accessed by users. The research model is shown in figure 1.

[Insert figure 1 about here]

An exploratory experiment was conducted to study the TWT of Web users under both *with* and *without* feedback conditions. Seventy subjects participated in the experiment. The subjects were undergraduate students enrolled in introductory MIS classes. The subjects had completed the development of their class home pages and were savvy Web users. The experiment was conducted as a laboratory assignment during class time in a University laboratory environment where high-speed Internet access (i.e. via T1 lines) was provided and expected by the students. The subjects were provided with a list of questions and were required to access specific Web pages to obtain the answers. All subjects used the same browser and interface (that was designed for the experiment). All of them received the same training session at the beginning of the experiment that familiarized them with the various buttons/icons available on the Web browser (specifically, the "STOP" button). The subjects began the task from a standard Web page that was designed specifically for the experiment. This standard Web page provided hyperlinks to the other Web pages that contained the information needed to obtain the answers to complete the assignment.

The subjects were randomly assigned into two groups for the experiment. One group of subjects (34 subjects) was provided with a feedback bar on their browser while the other group (36 subjects) was not. The feedback bar was a *moving* bar that signified to the users that the system was carrying out their request. It provided indications that the system was in a 'working' mode. The bar moved in a bi-directional manner (left to right, right to left, left to right, and so on) until the user's request was satisfied (i.e. the Web page was downloaded). Note that the feedback bar did not provide waiting duration information or status of the download *per se*, but simply an indication that the download was taking place.

More specifically, the subjects were asked to look up the names of 10 Web acceleration tools using the standard Web page provided to them. Of the 10 hyperlinks provided on the standard Web page, only 7 of them were working. Upon clicking on any of these 7 working hyperlinks, their corresponding Web page would appear instantaneously (i.e. with negligible download time). The fourth, seventh, and ninth hyperlinks triggered an *infinite* waiting time. For these 3 non-working hyperlinks, the subjects would have to click the "STOP" icon to terminate the wait. The subjects were required to click on all 10 hyperlinks which all of them did. Their TWT is the elapsed time between the moment the hyperlink was clicked (i.e. download request was made) and the moment the "STOP" button was clicked (i.e. download request was terminated). The computer log captured the elapsed times and all mouse-click actions for subsequent data analysis.

4. Research findings

As shown in table 3, the inclusion of a feedback bar significantly prolonged the waiting time of users. Thus, Hypothesis 1 is supported. Table 3 also shows that the *average* TWT for the *first* access to a *non-working hyperlink* was 13 seconds for the control group (no feedback bar) and 38 seconds for the treatment group (with feedback bar). The mode for TWT (i.e. where maximum number of abandonment occurred) was analyzed using intervals of one second. In other words, the mode refers to the time interval(s) where the maximum number of abandonment occurred. The mode for the first access to a non-working hyperlink for the control or 'no feedback bar' condition fell within the intervals of 5-6, 6-7 and 7-8 seconds, with a frequency (i.e. number of abandonment) of 4 in each interval. Hence, 33% (4x3/36) of the users terminated their first unsuccessful download request between 5-8 seconds. As for the 'feedback bar' condition, the mode occurred in various intervals between 15-46 seconds. The Mann-Whitney test indicates that the difference between the two conditions (with and without feedback bar) for the first non-working hyperlink is highly significant (as shown in table 3).

[Insert table 3 about here]

As subjects proceeded with the task, their TWT for accessing non-working hyperlinks decreased. This was probably because after encountering one unsuccessful download, the subjects no longer expected all the Web pages to be successfully downloaded. Hence, their expectations may have declined, causing their TWT to decrease. As shown in table 3, the *average* TWT for the first access to a second *non-working hyperlink* was 4 seconds for the control group (no feedback bar) and 17 seconds for the treatment group (with feedback bar). The Mann-Whitney test indicates that the difference is significant ($p<0.01$).

The mode for the first access to a second *non-working hyperlink* in the control group occurred in the intervals of 2-3 and 3-4 seconds, with a frequency (i.e. number of abandonment) of 11 in each interval (see table 4). Thus, 61% (11x2/36) of the users in the control condition gave up their wait to access the second non-working hyperlink between 2-4 seconds.

[Insert table 4 about here]

The mode for the first access to a second *non-working hyperlink* in the treatment group occurred in the interval of 2-3 seconds, with a frequency (i.e. number of abandonment) of 9 (see table 5). Thus, 26% (9/34) of the users in the treatment condition gave up their wait to access the second non-working hyperlink between the 2-3 second duration.

[Insert table 5 about here]

As shown in table 3, the *average* TWT for the first access to the last *non-working hyperlink* encountered was 3 seconds for the control group (no feedback bar) and 7 seconds for the treatment group (with feedback bar). The Mann-Whitney test indicates that the difference is significant ($p<0.01$). The mode for both groups was in the interval of 2-3 seconds, with a frequency (i.e. number of abandonment) of 13 (36%) for the control group (see table 6) and 8 (24%) for the treatment group (see table 7). Thus, the majority of the users gave up the wait between 2-3 seconds.

[Insert table 6 about here]

[Insert table 7 about here]

Overall, the results suggest that Web users' TWT peaks at approximately 2 seconds. This is in line with Shneiderman's (1986) and Miller's (1968) proposition that users are willing to wait for about 2 seconds before shifts in focus or interference with short-term memory occur.

4.1 Graphical illustrations of Web users' tolerable waiting time

Figures 2 and 3 show the distributions of TWT (in intervals of 5 seconds) for the first non-working hyperlink. Figures 4-7 show the distributions of TWT (in intervals of 5 seconds) for the other two non-working hyperlinks. The vertical axis, frequency, refers to the number of subjects who abandoned the wait during the time interval specified in the horizontal axis. This frequency is represented by the bar chart. The vertical axis also reflects the cumulative distribution (0-100%), which is presented by the line graph (or dots on the line graph).

[Insert figure 2 about here]

[Insert figure 3 about here]

As shown in figures 2 and 3, the subjects' waiting time was significantly prolonged when a feedback bar was provided on the Web browser (also see table 3). In the case where no feedback bar was provided, the mode for TWT was between 5-10 seconds (when analyzed in 5-second interval), as shown in figure 2.

As for the *first* accesses to *the other two non-working hyperlinks*, none of the users in the control setting (i.e. no feedback bar) waited more than 15 seconds, as shown in figures 4 and 5. This finding suggests that without indications from the system that it is working, Web users are not willing to wait for more than 15 seconds.

[Insert figure 4 about here]

[Insert figure 5 about here]

However, the scenario was different when a feedback bar was provided (see figures 6 and 7). When a feedback bar was provided, Web users' TWT increased. This finding is in line with that of Hui and Tse (1996) who found feedback information to increase users' sense of control and reduce users' uncertainty about the wait, which in turn increase their TWT.

[Insert figure 6 about here]

[Insert figure 7 about here]

5. Conclusion and future research

The availability of feedback prolongs Web users' TWT. Hence, it is beneficial for Web browsers or Web sites to provide feedback to users whenever there is an expected wait for page down download. Such information not only reduces uncertainty about the wait, but it also fills the time of the wait so Web users are less conscious about the duration of the wait.

Although TWT can vary under different circumstances and contexts, the findings from this study suggest that most users are willing to wait for only about two seconds for simple information retrieval tasks on the Web. This finding is consistent with most of the literature to date. Although this study was conducted in the Internet era, the findings on Web users' TWT are consistent with earlier research on (non-Internet-related) computer response time despite the different operating environments. According to Miller (1968), continuity of human thought processes is necessary for effective problem solving and a delay of more than 2 seconds may lead to psychological step-down discontinuities, which divert one's attention from the thought processes. The 2-second rule is also in line with Shneiderman's (1984) recommendation. With regard to research conducted in the Web context, the study by Galletta, Henry, McCoy and Polak (2002) provides some valuable insights on Web users' TWT. They examined delay times of 0, 2, 4, 6, 8, 10 and 12 seconds in an experiment and found that performance and behavioural intentions began to stabilize at 4-second delay. Hence, their results suggest that change in behavioural intentions takes place between the 2-4 second interval.

From this study, we found that Web users expect a response in about 2 seconds for simple information retrieval tasks on the Web. A 2-second response is needed to ensure 'smooth' interactions between the WWW and the users. The findings from this study also suggest that the upper bound for Web users' TWT is 15 seconds when the system does not provide any indication or feedback concerning the download (see figures 4 and 5). This is consistent with Nielsen's (1995) and Miller's (1968) prediction that response delays of approximately 15 seconds rule out conversational interaction between human and information systems. Miller (1968: 277) proposed that 'if response delays of more than 15 seconds will occur, the system had better be designed to free the user from physical and mental captivity, so that he can turn to other activities and get his displayed answer when it is *convenient to him* to do so.'

Future research will need to assess the applicability of the above findings to other tasks and contexts, and how different tasks and contexts might influence the TWT of Web users. Interestingly, in contrast to general expectations, Selvidge (2003) found that TWT is not affected by task type (information retrieval, online purchasing, downloading a text file). Regardless, the effect of task nature (e.g. netsurfing, browsing, querying) on TWT is one area that needs further research. To stimulate future research, the following sub-sections provide further discussions on the various variables that may affect Web users' TWT.

5.1 Nature of task and waiting time

Netsurfing refers to the scenario where users explore and wander around various websites without any clear objectives or purposes (Hayes 1995). In querying and browsing, the users possess a purpose. Compared to querying, browsing is a more casual search approach and is often practiced during activities such as exploratory learning. Querying refers to serious search with specific requirement and often involves a highly complicated search strategy such as the use of Boolean operators (i.e. AND/OR). People who surf or browse the net are generally not willing to spend the same amount of time and resources as they would if they were querying for specific information (Reaux and Carroll 1997). Hence, the nature of the task is expected to influence users' tolerance on download waiting time. For example, Rose and Straub (2001), and Rose, Lees and Meuter (2001) have studied download time in a completely different context from this study (i.e. the e-commerce context) by examining its impact on consumer attitude toward e-service retailer and patronage intentions toward e-retailers.

5.2 Waiting duration information and waiting time

Another interesting area for research is to study the effect of waiting duration information on TWT. Providing users with waiting duration information is a common practice available on most Web browsers. For example, while a page is being downloaded, Netscape Navigator and Microsoft Explorer provide status or retrieval information on the bar located at the bottom of the browser window. The effect of this type of information, though believed to make long waiting time more tolerable, needs to be empirically studied and verified. For example, the study by Hui and Zhou (1996) shows that providing customers with waiting duration information does not reduce their perceived waiting duration, which is contrary to common beliefs. They also show that the status duration information increases TWT not by changing customers' perception of the waiting duration but by increasing their perceived cognitive control (Folkman 1984). Hence, the mediating effects of providing status duration information will need to be better understood and further investigated.

5.3 Display techniques and waiting time

One technique that has been used to ease users' frustration when waiting for Web pages to appear is the interlacing technique. Using this technique, the image first appears as a vague image, and then slowly clears up and becomes more focused as more data is received. How does this technique compare to the usual top-down approach? Intuitively, one would expect the interlacing technique to be superior. Allan (1979), however, indicates that, 'a filled interval is judged as longer than an empty interval of the same stimulus duration'. This result is surprising and it seems to contradict our findings. If Allan's finding can be applied to the interlacing technique, then interfacing would result in longer perceived waiting time. More empirical research is, therefore, needed to compare the different strategies of displaying Web pages and their impact on waiting time, as well as to study the effect of filled intervals of different kinds.

5.4 Relationship between other factors and waiting time

Many other factors can affect Web users' TWT including the amount of multi-media or graphics available on the Web site/page, users' expectations of download time (e.g. dial-up versus high-speed Web access), users' goals, incentives or rewards for completion of the task, demographics of users (e.g. experience, age, gender, personality, culture), availability of alternative Web pages, time pressure, and environmental factors. These factors may affect not only Web users' TWT, but also their perceptions, attitudes, intentions, and performance. A recent study has demonstrated that culture affects perceived delay time and attitude toward download delay (Rose *et al*. 2003). The types of media available on a Web site also affect the users' willingness to wait for download (Jacko *et al*. 2000). Future research is needed to examine the above factors and their effect on Web users' perceptions, attitudes, intentions, behaviour, and performance. In other words, to develop a comprehensive theory on the factors influencing Web users' TWT, it is important to understand the relationships between the various possible dependent variables, such as perceptions of waiting time, attitudes toward the wait (e.g. satisfaction and frustration), intentions to visit the Web page/site again or to give up the wait, actual behaviour to abandon the wait or to visit alternate Web page/site, and user performance on the task.

5.5 Implications for practice and research

According to Zona Research (2001: 6), 'for every second of latency over normal expectations of that page, a Web transaction accumulates a demerit'. In the same report, the risk of losing revenue due to site abandonment was also discussed. Designers and operators of websites need to ensure that their sites can be accessed within a reasonable amount of time, i.e. within 2 seconds for every page, or they risk losing revenues. The 2-second rule can also be used by network service providers and administrators to determine their hardware and software requirements and investments, and by Web designers to decide on the optimal design of websites. For example, unnecessary graphics and multimedia should be avoided. Web acceleration tools may be installed to boost up the download speed, for example by loading Web pages and graphics in advance (i.e. based on predictions of next Web page access).

In this research, the concept of perceived waiting time was not studied. Research has suggested that there is a linear relationship between perceived time and actual time (Rule *et al*. 1970, Allan 1979, Hornik 1984). In fact, Hornik (1984) found that individuals tend to overestimate waiting time. This is consistent with Cottle's (1976) research where it was found that subjects have a tendency to overestimate passive durations (such as waiting) and underestimate active durations of time. Antonides *et al*. (2002), however, found a non-linear relationship between perceived and actual waiting time. They also found that information about the expected waiting time significantly reduced the overestimation of waiting time, although it increased the negative effect of perceived waiting time on wait evaluations.

It is important to study perceived waiting time in future research because it could directly influence a user's decision to give up or continue waiting for the download of a Web page (Weinberg, 2000). In fact, perceived waiting time may be more relevant and important than true waiting time, as a user seldom bases his/her decision to continue or quit waiting by the actual length of time s/he has waited, but rather by the amount of time s/he is perceived to have waited. Consequently, different types of techniques can be used to reduce Web users' perceived waiting time which may prolong users' actual waiting time. For example, Allan (1979) pointed out that perceived duration could be influenced by non-temporal characteristics such as modality, nature (filled vs. empty), energy, and complexity.

Given that long waiting time has always been one of the leading concerns for Web users (Lightner *et al*. 1996, GVU 1998, Selvidge 1999, 2003), it is important for researchers and practitioners to: 1) understand users' waiting behaviour in accessing the Web, 2) propose and evaluate techniques to reduce users' actual and perceived waiting time, and 3) provide guidelines that take into account the trade-offs between download/access time and aesthetics of Web pages. Finally, it is hoped that this research will stimulate the interest of other researchers to examine issues related to TWT for different

Web activities in different contexts. Researchers and practitioners can contribute to this area of research by proposing mechanisms to either reduce users' waiting time or make their Web experience more pleasing or tolerable, as well as evaluating the effectiveness and impact of these mechanisms on users' perceptions and behaviour.

References

Allan, L.G., 1979, The perception of time. *Perception and Psychophysics*, **26**(5), 340-354.

Antonides, G., Verhoef, P.C., van Aalst, M., 2002, Consumer perception and evaluation of waiting time: a field experiment. *Journal of Consumer Psychology*, **12**(3), 193-202.

Bailey, B., 2001, Acceptable computer response times. *UI Design Update Newsletter*, April 2001. Available at: http://www.humanfactors.com/downloads/apr012.htm#bobbailey.

Chen, B., 2002, The 30-second rule. *Network World*, 22 July 2002. Available at: http://www.nwfusion.com/columnists/2002/0722chen.html.

Cottle, T.J., 1976, *Perceiving Time: A Psychological Investigation with Men and Women* (New York: Wiley).

Dellaert, B.G.C. and Kahn, B.E., 1999, How tolerable is delay? Consumers' evaluations of Internet Web sites after waiting. *Journal of Interactive Marketing*, **13**(1), 41-54.

Folkman, S., 1984, Personal control and stress and coping process: a theoretical analysis. *Journal of Personality and Social Psycholog*y, **46**, 839-852.

Galletta, D., Henry, R., McCoy, S. and Polak, P., 2002, Web site delays: how slow can you go?" Presented at the *First Annual Pre-ICIS Workshop on HCI Research in MIS*, 14 December 2002 (Barcelona, Spain: AIS Special Interest Group on Human Computer Interaction), forthcoming in *Journal of Association for Information Systems*.

Gilliland, A.R., Hofeld, J. and Eckstrand, G., 1946, Studies in time perception. *Psychological Bulletin*, **43**, 162-176.

Graphics, Visualization and Usability (GVU) Centre, 1998, GVU's user surveys. Georgia Tech Research Corporation, October 1998, available at: http://WWW.cc.gatech.edu/gvu/user_ surveys/.

Guynes, J.L., 1988, Impact of system response time on state anxiety. *Communications of the ACM*, **31**(3), 342-347.

Hayes, M., 1995, Working online or wasting time? *Information Week*, 1 May 1995, 38-51.

Hornik, J., 1986, Subjective vs. objective time measures: a note on the perception of time in consumer behavior. *Journal of Consumer Research*, **11**, 615-618.

Hoxmeier, J.A. and DiCesare, C., 2000, System response time and user satisfaction: an experimental study of browser-based applications. *Proceedings of the Americas Conference on Information Systems*, 10-13 August 2000 (Long Beach, California: Association for Information Systems), pp. 140-145.

Hui, M.K. and Tse, D., 1996, What to tell consumers in waits of different lengths: an integrative model of service evaluation. *Journal of Marketing*, **60**(2), 81-90.

Hui, M.K. and Zhou, L., 1996, How does waiting duration information influence customers' reactions to waiting for services? *Journal of Applied Social Psychology*, **26**(19), 1702-1717.

Jacko, J.A., Sears, A. and 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.

Katz, K., Larson, B. and Larson, R., 1991, Prescription for the waiting in line blues: entertain, enlighten and engage. *Sloan Management Review*, **32**(2), 44-53.

Lightner, N.J., Bose, I. and Salvendy, G., 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 transaction. *Proceedings of AFIPS Fall Joint Computer Conference*, **33**, 267-277.

Nielsen, J., 1993, Response times: the three important limits. Available at http://www.useit.com/papers/responsetime.html. Excerpt from Chapter 5 of *Usability Engineering* by Jakob Nielsen, Academic Press, 1993.

Nielsen, J., 1995, Guidelines for multimedia on the web. *Jakob Nielsen's Alertbox for December, 1995*. Available at http://www.useit.com/alertbox/9512.html.

Nielsen, J., 1996, Top ten mistakes in Web design. *Jakob Nielsen's Alertbox for May, 1996*. Available at http://www.useit.com/alertbox/9605.html.

Nielsen, J., 1997, The need for speed. *Jakob Nielsen's Alertbox for March, 1997*. Available at http://www.useit.com/alertbox/9703a.html.

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

Nielsen, J., 2000, *Designing Web Usability: the practice of simplicity* (Indianapolis: New Riders).

Osuna, E.E., 1985, The psychological cost of waiting. *Journal of Mathematical Psychology*, **29**(1), 82-105.

Pitkow, J.E. and Kehoe, C.M., 1996, Emerging trends in the WWW user population. *Communications of the ACM*, **39**(6), 106-108.

Ramsay, J., Barbesi, A. and Preece, J., 1998, A psychological investigation of long retrieval times on the world wide web. *Interacting with Computers*, **10**(1), 77-86.

Reaux, R.A. and Carroll, J.M., 1997, Human factors in information access of distributed systems. In *Handbook of Human Factors & Ergonomics*, edited by Gavriel Salvendy (New York: Wiley), Second Edition.

Rose, G.M., Evaristo, R. and Straub, D., 2003, Culture and consumer responses to web download time: a four-continent study of mono and polychronism. *IEEE Transactions on Engineering Management*, **50**(1), 31-44.

Rose, G.M., Lees, J. and Meuter, M.L., 2001, A refined view of download time impacts on e-consumer attitudes and patronage intentions toward e-retailers. *International Journal of Media Management*, **3**, 105-111.

Rose, G.M. and Straub, D.W., 2001, The effect of download time on consumer attitude toward the e-service retailer. *e-Service Journal*, **1**, 55-76.

Rule, S.J., Curtis, D.W. and Markley, R.P., 1970, Input and output transformations from magnitude estimation. *Journal of Experimental Psychology*, **86**(1), 343-349.

Selvidge, P., 1999, How long is *too long* for a website to load? *Usability News*, **1**(2). Available at: http://psychology.wichita.edu/surl/usabilitynews/1s/time_delay.htm.

Selvidge, P., 2003, Examining tolerance for online delays. *Usability News*, **5**(1). Available at: http://psychology.wichita.edu/surl/usabilitynews/51/delaytime.htm.

Shneiderman, B., 1984, Response time and display rate in human performance with computers. *Computing Surveys*, **16**, 265-285.

Taylor, S., 1994, Waiting for service: the relationship between delays and evaluations of service. *Journal of Marketing*, **58**(2), 56-69.

Weinberg, B.D., 2000, Don't keep your Internet customers waiting too long at the (virtual) front door. *Journal of Interactive Marketing*, **14**(10), 30-38.

Zakay, D. and Hornik, J., 1991, How much time did you wait in line? A time perception perspective. In *Time and Consumer Behaviour*, edited by Jean-Charles Chebat and V. Venkatesan (Montreal: University of Quebec at Montreal).

Zona Research Report, 1999, The need for speed, July 1999.

Zona Research, 2001, The need for speed II. *Zona Market Bulletin*, **5**, April 2001. Available at: http://www.keynote.com/downloads/Zona_Need_For_Speed.pdf.

Table 1. Summary of users' tolerable waiting time for computer response

Study	Findings/Recommendations
Miller (1968)	• Delay of 2 seconds is the limit before interference with short-term memory occurs
Nielsen (1993, 1995, 1996)	• Delay of 0.1 second is perceived as instantaneous access • Delay of 1.0 second is the limit for users' flow of thought to stay uninterrupted • Delay of 10 seconds is the limit for keeping users' attention/focus on the dialogue
Shneiderman (1984)	• Delay of 2 seconds is the limit where response to simple commands becomes unacceptable to users

Table 2. Summary of users' tolerable waiting time for Web page download

Study	Findings/Recommendations
Ramsey, Barbesi and Preece (1998)	• Delay of 41 seconds is suggested as the cut-off for long delays based on users' perceptions
Selvidge (1999)	• Delay of 30 seconds is suggested as the cut-off based on users' performance and frustration levels
Nielsen (1993, 1995, 1996)	• Delay of 15 seconds is tolerable in the Web context
Hoxmeier and DiCesare (2000)	• Delay of 12 seconds causes satisfaction to decrease
Galletta, Henry, McCoy and Polak (2002)	• Delay of 4 seconds causes performance and behavioural intentions to stabilize whereas attitudes remain unchanged after delay exceeds 8 seconds

Table 3. Statistics on waiting time for WWW access

	Subjects' waiting time for first access to non-working hyperlinks		
	1st non-working hyperlink	2nd non-working hyperlink	3rd non-working hyperlink
Control (no FB) (36 subjects)	**Mean = 13 sec.** **Median = 9 sec.** **Mode = 5-8 sec.**	**Mean = 4 sec.** **Median = 3.6 sec.** **Mode = 2-4 sec.**	**Mean = 3.3 sec.** **Median = 2.5 sec.** **Mode = 2-3 sec.**
Treatment (with FB) (34 subjects)	**Mean = 37.6 sec.** **Median = 22.6 sec.** **Mode = 15-16, 20-22, 45-46 sec.**	**Mean = 17 sec.** **Median = 8.4 sec** **Mode = 2-3 sec**	**Mean = 6.7 sec.** **Median = 4.3 sec** **Mode = 2-3 sec**
Mann-Whitney test	p=0.000	p=0.002	p=0.004

* FB = feedback during Web page download

Table 4. TWT for first access to second non-working hyperlink in control condition

Control	TWT ≤ 1 sec.	1 sec. <TWT≤ 2 sec.	2 sec. <TWT≤ 3 sec.	3 sec. <TWT≤ 4 sec.	4 sec. <TWT≤ 5 sec.	5 sec. <TWT≤ 6 sec.	6 sec. <TWT≤ 7 sec.
Frequency	0	3	**11**	**11**	3	1	2
Percentage	0%	8% (3/36)	**31% (11/36)**	**31% (11/36)**	8% (3/36)	3% (1/36)	6% (2/36)

Table 5. TWT for first access to second non-working hyperlink in treatment condition

Treatment	TWT ≤ 1 sec.	1 sec. <TWT≤ 2 sec.	2 sec. <TWT≤ 3 sec.	3 sec. <TWT≤ 4 sec.	4 sec. <TWT≤ 5 sec.	5 sec. <TWT≤ 6 sec.	6 sec. <TWT≤ 7 sec.
Frequency	0	0	**9**	4	1	1	1
Percentage	0%	0%	**26% (9/34)**	12% (4/34)	3% (1/34)	3% (1/34)	3% (1/34)

Table 6. TWT for first access to last/third non-working hyperlink in control condition

Control	TWT ≤ 1 sec.	1 sec. <TWT≤ 2 sec.	2 sec. <TWT≤ 3 sec.	3 sec. <TWT≤ 4 sec.	4 sec. <TWT≤ 5 sec.	5 sec. <TWT≤ 6 sec.	6 sec. <TWT≤ 7 sec.
Frequency	0	8	**13**	6	2	4	1
Percentage	0%	22% (8/36)	**36% (13/36)**	17% (6/36)	6% (2/36)	11% (4/36)	3% (1/36)

Table 7. TWT for first access to last/third non-working hyperlink in treatment condition

Treatment	TWT ≤ 1 sec.	1 sec. <TWT≤ 2 sec.	2 sec. <TWT≤ 3 sec.	3 sec. <TWT≤ 4 sec.	4 sec. <TWT≤ 5 sec.	5 sec. <TWT≤ 6 sec.	6 sec. <TWT≤ 7 sec.
Frequency	0	4	**8**	4	4	0	0
Percentage	0%	12% (4/34)	**24% (8/34)**	12% (4/34)	12% (4/34)	0%	0%

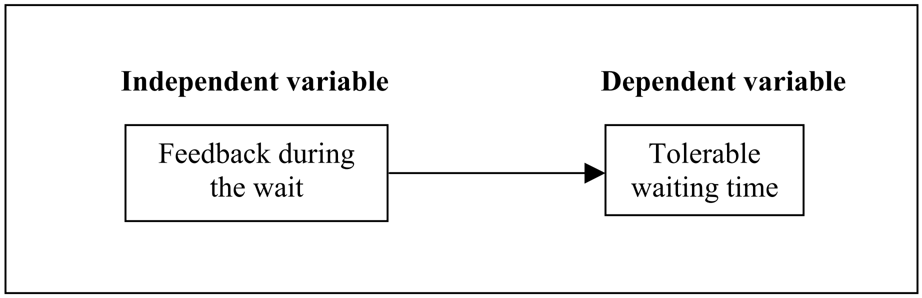

Figure 1: Research model

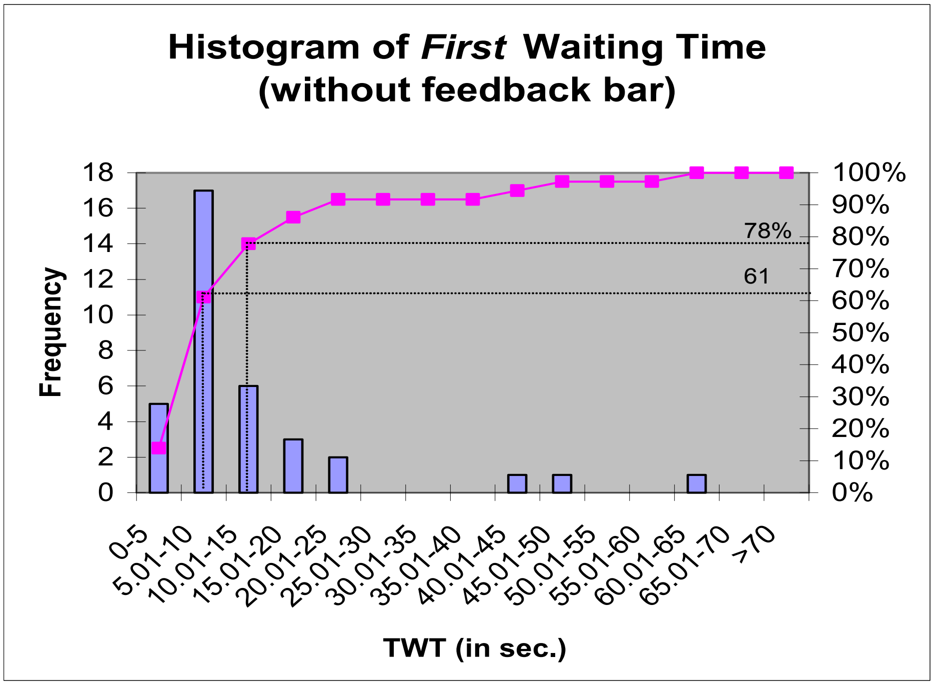

Figure 2. Cumulative distribution of TWT in the absence of a feedback bar

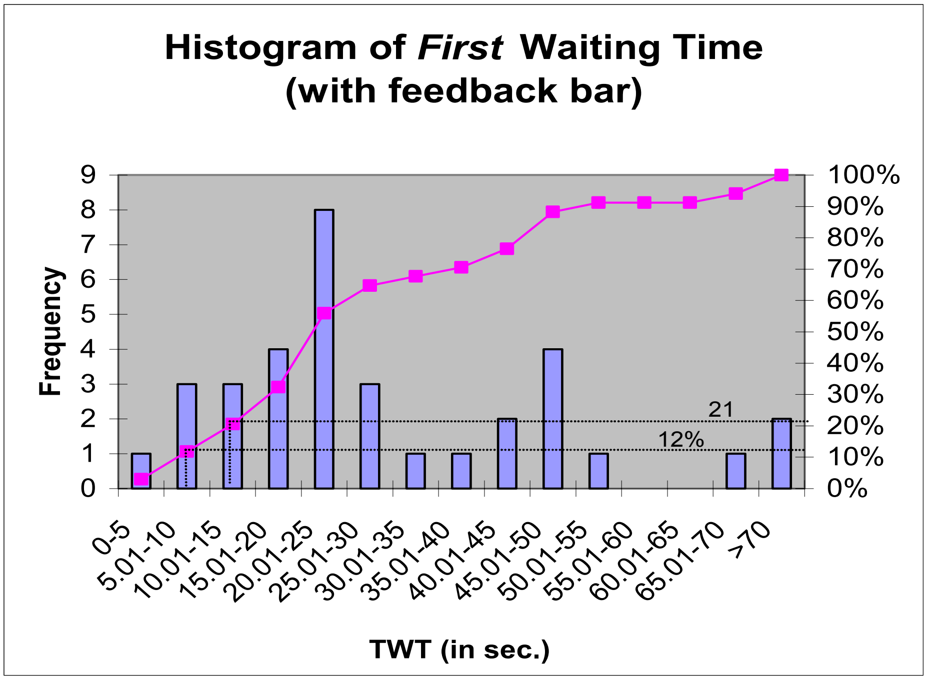

Figure 3. Cumulative distribution of TWT in the presence of a feedback bar

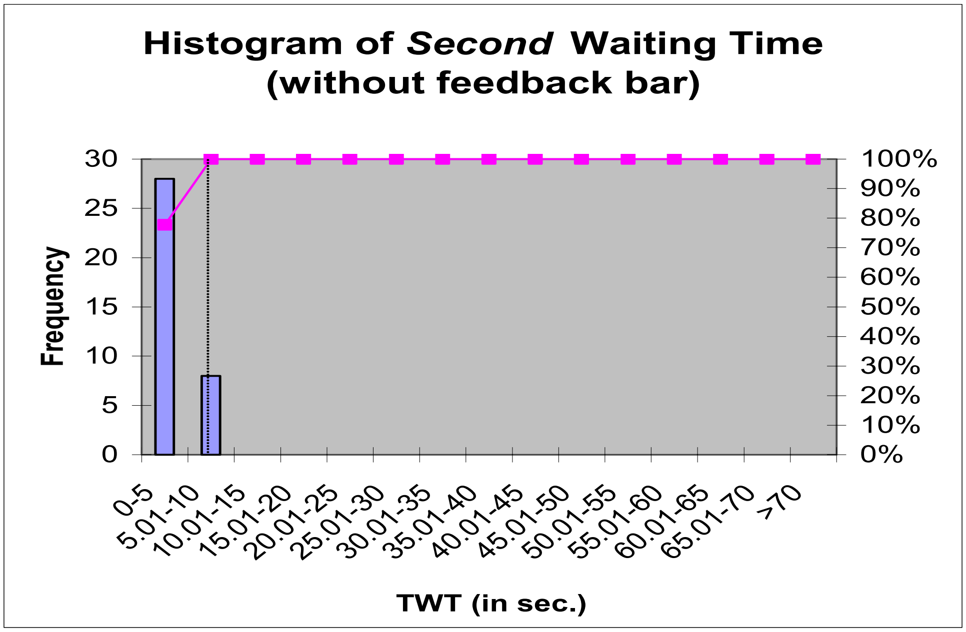

Figure 4. Cumulative distribution of TWT for first access to a second non-working hyperlink in the absence of a feedback bar

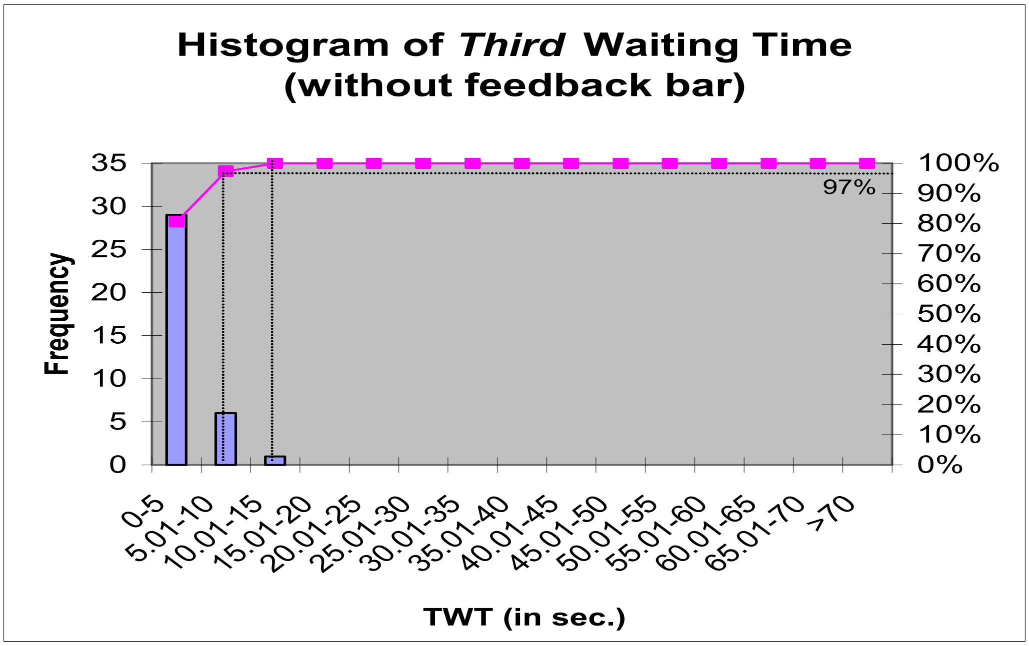

Figure 5. Cumulative distribution of TWT for first access to the final/third non-working hyperlink in the absence of a feedback bar

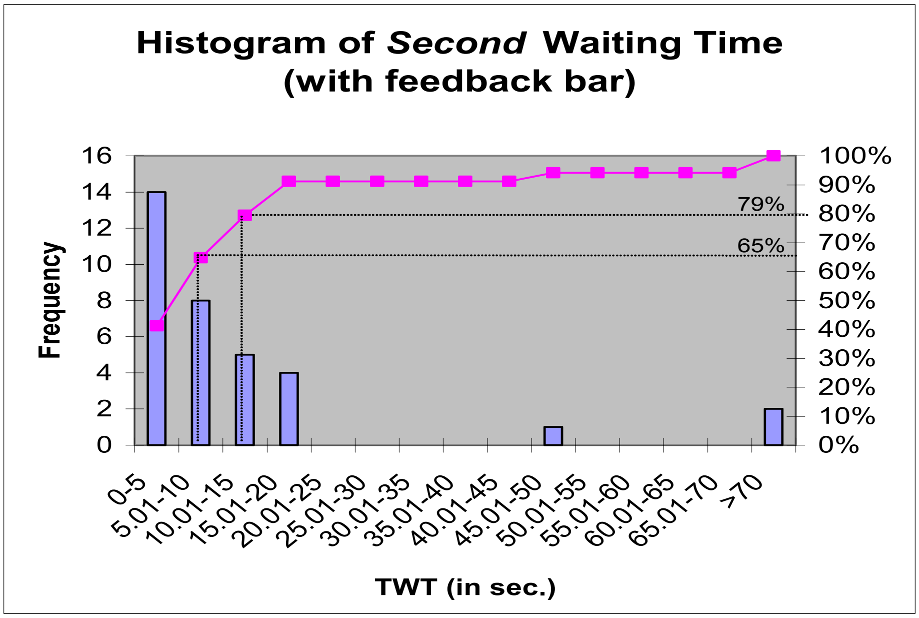

Figure 6. Cumulative distribution of TWT for first access to a second non-working hyperlink in the presence of a feedback bar

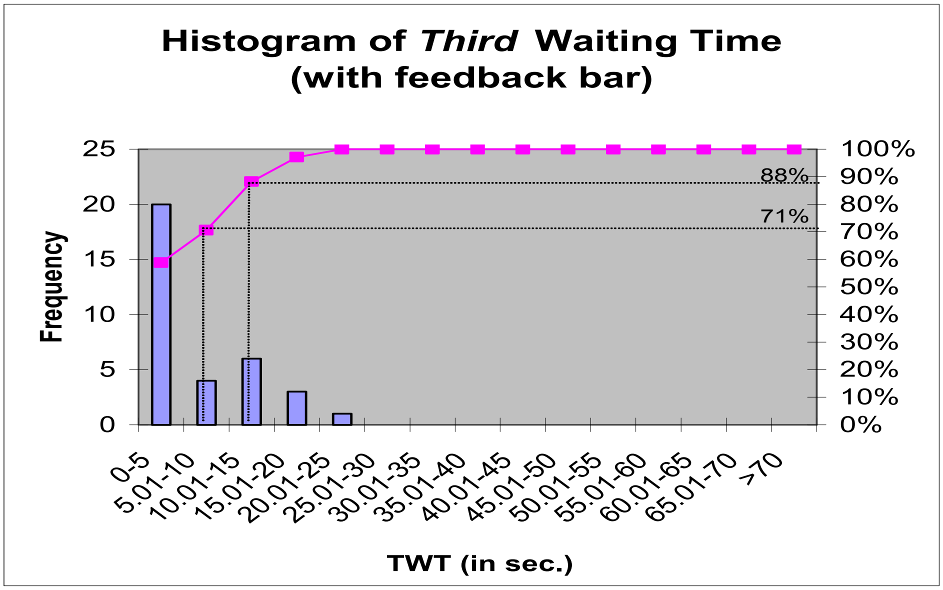

Figure 7. Cumulative distribution of TWT for first access to the final/third non-working hyperlink in the presence of a feedback bar