

**EFFECTS OF INTERFACE FORMAT, FEEDBACK STYLE, AND
SYSTEM LAG ON THE USABILITY OF
HAND-HELD INTERNET CONTROLLERS**

Feyza F. Ergen

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Industrial and Systems Engineering

Robert J. Beaton, Chair
John Deighan
Willard Farley

December 17, 1996
Blacksburg, Virginia

Keywords: User-interface design, human-computer interaction, system response time, feedback, tree-menu structures, Internet

Copyright 1997, Feyza F. Ergen

EFFECTS OF INTERFACE FORMAT, FEEDBACK STYLE, AND SYSTEM LAG ON THE USABILITY OF HAND-HELD INTERNET CONTROLLERS

Feyza F. Ergen

(ABSTRACT)

The increasing popularity of the World Wide Web (WWW) has created a new market: Web access through television to accommodate those who either cannot afford existing hardware or are intimidated by computers. Current efforts to combine the WWW and television have targeted potential novice users. One of the approaches for creating a WWW browsing system that is both simple to use and inexpensive is the utilization of the existing cable system to provide Web access through television. Some unique characteristics of this browsing system are fast access to the Web, the use of nine buttons on a universal remote control, and an index structure for reaching Web sites. Since browsing the Web through television is relatively new, many interface issues need to be examined. The purpose of this research was to investigate potential user interface designs for this WWW browser and to evaluate the usability of the nine-button interface.

Sixteen volunteers participated in the experiment and were asked to navigate to specific Web sites with two interface formats, five different system lag times (0.2, 0.7, 1.3, 2.0, and 3.0 seconds), and three feedback styles (active feedback, passive feedback and no feedback). Participants were prescreened for their experience with computers and browsing the Web. The experiment was conducted in a living room setting to simulate real life situations and participants were given a total of 42 tasks to complete throughout the experiment. Each task consisted of navigating through the tree structure with either one of the two interface formats until reaching a designated Web site. The number of errors committed and task completion times were recorded. In addition, participants were asked to rate the WWW browser system after each task as well as after the entire experiment.

Participants preferred the 0.2 second system lag and the active feedback style. Overall, they committed fewer errors and took less time to complete tasks with the tabbing interface than with the one-to-one mapping interface. Experienced participants committed more errors than did the inexperienced ones. Increase of system lag time was determined to have a greater adverse effect on the tabbing interface than on the one-to-one mapping interface.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Robert Beaton for his guidance, encouragement, and support throughout this research. My other committee members, John Deighan and Willard Farley, are also acknowledged for their help and advice.

Special thanks to John Deighan for the time he invested in developing the Director prototype. Thanks to Will Vest for assembling the receiver and to Larry Barfood for voluntarily taking screen shots and making my life easier.

This research was funded by Fernando Morales through the Center for Wireless Telecommunications. Their support is greatly appreciated.

Special thanks to my husband, Tsunou, for his moral support and his valuable input to this research. Finally, I would like to thank my parents and my brother for their love, understanding, and unconditional support.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
LIST OF FIGURES.....	vi
LIST OF TABLES	viii
CHAPTER 1. INTRODUCTION.....	1
CHAPTER 2. LITERATURE REVIEW.....	3
SYSTEM LAG.....	3
FEEDBACK.....	6
INTERFACE FORMAT.....	7
CHAPTER 3. EXPERIMENTAL METHOD	9
FACILITIES AND APPARATUS.....	9
PARTICIPANTS	9
EXPERIMENTAL DESIGN.....	9
<i>Independent Variables</i>	12
Interface Format.....	12
Feedback Style.....	15
System Lag.....	15
<i>Dependent Variables</i>	15
<i>Mathematical Model</i>	22
EXPERIMENTAL PROCEDURE.....	25
<i>Pre-experimental Training</i>	25
<i>Experimental Session</i>	25
DATA ANALYSIS	25
CHAPTER 4: RESULTS.....	26
TASK COMPLETION TIME.....	26
NUMBER OF ERRORS.....	26
SUBJECTIVE RATINGS	26
<i>System Lag Time</i>	26
<i>Task Difficulty</i>	34
<i>Feedback Style</i>	40
<i>Usability</i>	45
<i>Subjective ratings for the two Interface Formats</i>	46
<i>Participants' preferences</i>	46
CHAPTER 5: DISCUSSION.....	55
SYSTEM LAG.....	55
FEEDBACK STYLE	55
INTERFACE FORMAT.....	55
EXPERIENCE LEVEL	56
CHAPTER 6: CONCLUSION.....	57

REFERENCES.....	58
APPENDIX A: Prescreening Questions.....	60
APPENDIX B: Tasks Performed during the Experiment.....	61
APPENDIX C: Subjective Evaluation Questions	64
VITA	67

LIST OF FIGURES

Figure 1.1 Illustration of how the WWW is linked to television through the existing cable system by using a 900 MHz wireless technology.....	2
Figure 2.1 User think time as a function of computer response time in four different types of subject testings.....	4
Figure 2.2 Error rates as a function of response time for a complex task.....	4
Figure 3.1 Setup of the experiment room.....	10
Figure 3.2 The universal remote control used in the experiment.....	11
Figure 3.3 Setup for the experiment.....	11
Figure 3.4 Experimental design.....	12
Figure 3.5 The two interface stages.....	13
Figure 3.6 Illustration of how the nine buttons on the remote control work with the two interface formats.....	16
Figure 3.7 Step-by-step procedure for reaching the Boeing web site using the one-to-one mapping interface format.....	17
Figure 3.8 Cancel screen in the one-to-one mapping interface.....	18
Figure 3.9 Functions of the buttons on the remote control.....	18
Figure 3.10 Step-by-step procedure for reaching the Boeing web site using the tabbing interface format.....	19
Figure 3.11 Screens that appeared (a) when participants chose a topic that did not lead to web pages and (b) when participants pressed “quit”.....	20
Figure 3.12 Three different types of feedback styles.....	21
Figure 3.13 Timeline for the two interface formats.....	23
Figure 3.14 Average task completion time as a function of (a) System lag time and (b) Network download time.....	24
Figure 4.1 Effect of Interface Format on task completion time.....	28
Figure 4.2 Effect of System Lag on task completion time.....	28
Figure 4.3 Effect of Experience Level on number of errors committed.....	30
Figure 4.4 Effect of Interface Format on number of errors committed.....	30
Figure 4.5 Effect of Interface Format on subjective ratings for System Lag.....	32
Figure 4.6 Effect of System Lag on subjective ratings for System Lag.....	32
Figure 4.7 Effect of Interface Format × Feedback Style interaction on subjective ratings for System Lag.....	33
Figure 4.8 Effect of Experience Level on subjective ratings for task difficulty.....	36
Figure 4.9 Effect of Interface Format on subjective ratings for task difficulty.....	36
Figure 4.10 Effect of Interface Format × Feedback Style interaction on subjective ratings for task difficulty.....	37

Figure 4.11 Effect of Interface Format × System Lag interaction on subjective ratings for task difficulty.....	38
Figure 4.12 Effect of Interface Format × Feedback Style × System Lag interaction on subjective ratings for task difficulty.....	39
Figure 4.13 Effect of Interface Format on subjective ratings for Feedback Style.....	42
Figure 4.14 Effect of Feedback Style on subjective ratings for Feedback Style.....	42
Figure 4.15 Effect of Interface Format × Feedback Style interaction on subjective ratings for Feedback Style.....	43
Figure 4.16 Effect of Interface × System Lag interaction on subjective ratings for Feedback Style.....	44
Figure 4.17 Effect of Experience Level on subjective ratings for usability of the system.....	48
Figure 4.18 Effect of Feedback Style on subjective ratings for usability of the system.....	48
Figure 4.19 Effect of System Lag on subjective ratings for usability of the system.....	49
Figure 4.20 Effect of Interface Format × Feedback Style interaction on subjective ratings for usability of the system.....	50
Figure 4.21 Effect of Interface Format × System Lag interaction on subjective ratings for usability of the system.....	51
Figure 4.22 Effect of Interface Format × Feedback Style × System Lag interaction on subjective ratings for usability of the system.....	52
Figure 4.23 Results of the subjective ratings for each Interface Format.....	54

LIST OF TABLES

Table 4.1 ANOVA Summary Table for Average Task Completion Time.....	27
Table 4.2 Newman-Keuls Comparison of Task Completion Time among System Lag Times.....	29
Table 4.3 ANOVA Summary Table for Average Number of Errors.....	29
Table 4.4 ANOVA Summary Table for Subjective Ratings on System Lag.....	31
Table 4.5 Newman-Keuls comparison of System Lag subjective ratings among System Lag times.....	33
Table 4.6 Newman-Keuls comparison of System Lag subjective ratings for Interface Format × Feedback Style interaction.....	33
Table 4.7 ANOVA Summary Table for Subjective Ratings on Task Difficulty.....	35
Table 4.8 Newman-Keuls comparison of task difficulty subjective ratings for Interface Format × Feedback Style interaction.....	37
Table 4.9 Newman-Keuls comparison of task difficulty subjective ratings for Interface Format × System Lag interaction.....	38
Table 4.10 Newman-Keuls Comparison of Task Difficulty Subjective Ratings for Interface Format × Feedback Style × System Lag Interaction.....	40
Table 4.11 ANOVA Summary Table for Subjective Ratings on Feedback Style.....	41
Table 4.12 Newman-Keuls comparison of Feedback Style subjective ratings among Feedback Styles.....	43
Table 4.13 Newman-Keuls comparison of Feedback Style subjective ratings or Interface Format × Feedback Style interaction.....	44
Table 4.14 Newman-Keuls comparison of Feedback Style subjective ratings for Interface Format × System Lag Interaction.....	45
Table 4.15 ANOVA Summary Table for Subjective Ratings on the Usability of the System.....	47
Table 4.16 Newman-Keuls comparison of usability subjective ratings among Feedback Styles.....	49
Table 4.17 Newman-Keuls comparison of usability subjective ratings among System Lag times.....	49
Table 4.18 Newman-Keuls comparison of usability subjective ratings for Interface Format × Feedback Style interaction.....	50
Table 4.19 Newman-Keuls comparison of usability subjective ratings for Interface Format × System Lag interaction.....	51
Table 4.20 Newman-Keuls comparison of usability subjective ratings for Interface Format × Feedback Style × System Lag interaction.....	53

CHAPTER 1. INTRODUCTION

The creation of the World Wide Web (WWW) is revolutionizing the way people receive, send, and process information. Users can process information from all around the world anytime of the day quickly and inexpensively.

The WWW is handled through the Internet, which requires the use of a computer and some type of communication line such as the modem or the ethernet; the setup can be overwhelming for many potential new users who wish to access the WWW but either cannot afford the system or are intimidated by computers. Current efforts are to simplify the WWW access and to make it available to these potential new users. One approach is to combine the WWW and the television to create a browsing system that is both simple to use and inexpensive. Today many major electronics companies are marketing products that make the WWW accessible through television and many other companies are expected to follow. One approach currently being studied is linking the WWW to the television through the existing cable system by using a 900 MHz wireless technology.

This new technology takes advantage of the Internet in order to make interactive television a reality. It operates at ethernet speeds, making it uniquely fast compared to other existing systems. It enables users to have the comfort and convenience of performing a number of operations through their television sets such as accessing the WWW and ordering a product as they are watching a commercial.

The browsing system consists of a transmitter, a 900 MHz cell repeater, and the host computer system. Since the system has many uses, the transmitter has different modes of operation such as audio input, information transmission, WWW browser, and serial data input. This study is limited to the “WWW browser” mode of the system.

There will be a specific cable television channel assigned for the WWW Browser. Once the users tune their television sets to the “Web” channel, they will be able to browse the WWW by using their universal remote control. Only nine of the buttons on the remote control will be used. Since it will be possible to use any universal remote control and there are many modes of operation for the system, it will not be possible to label the nine buttons specifically for the WWW browsing function.

The universal remote control communicates with the transmitter that transfers users’ requests to the cell repeater via a 900 MHz radio channel; the cell repeater in turn sends these messages to the host system. The communication between the cell repeater and the host system can be carried by a standard phone line, Cellular Digital Packet Data (CDPD), ethernet

connection, or cable modem. The host system processes these requests and takes appropriate action (Figure 1.1). Only one of the users assigned to a particular cell repeater can browse the WWW at a time. However, there may be many cell repeaters for each host system.

The purpose of this study was to develop the user-interface design for the WWW browser, to evaluate the usability of the nine-button interface, and to investigate the effects of interface format, system lag, and feedback style on the usability of the system.

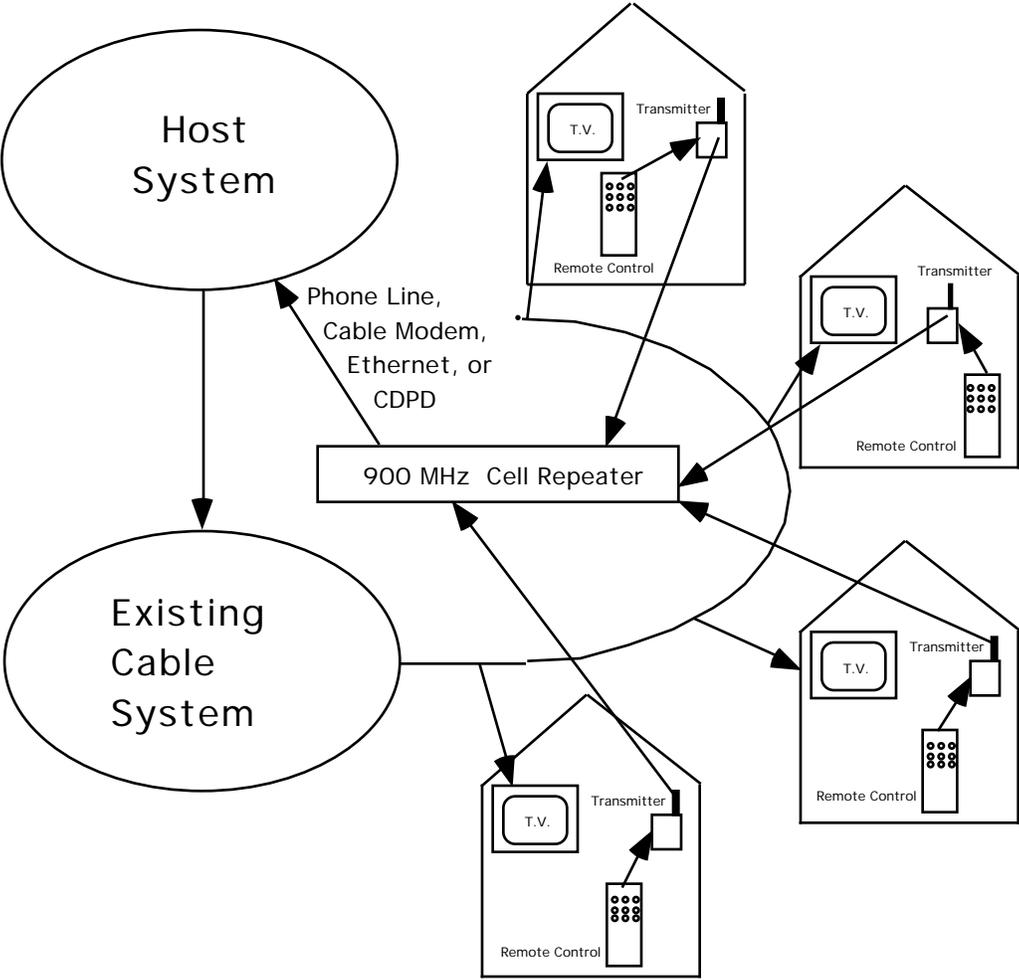


Figure 1.1 Illustration of how the WWW is linked to television through the existing cable system by using a 900 MHz wireless technology.

CHAPTER 2. LITERATURE REVIEW

The three factors that are considered in this study are system lag, feedback style, and interface format. The Experimental Method section describes how the effects of these variables on the usability of the system are determined. In this section some of the literature on these three topics is summarized.

System Lag

System lag is defined as the time from the moment an input is entered to the moment the system responds to the user. This is similar to the system response time (SRT) which is defined as the time between the user's input and the system's response (Thum, et. al. 1995).

SRT and system breakdowns have been determined to cause stress and reduce productivity (Kuhmann, 1989; Shneiderman, 1984). Traditionally it was thought that the shorter the SRTs the better; however, various studies have shown that desired SRTs depend on a number of factors such as work style, amount of experience of the user, the incentive, and the performed task.

Prolonged SRTs are known to cause involuntary working pauses. Researchers have shown that long SRTs are not used for relaxation or as user thought time, but they generate psycho-physiological stress (Kuhmann, 1989). Short SRTs, however, cause users to increase their processing speed forcing an agitated state of mind which causes many computing errors accompanied by stress (Kuhmann, 1989). Researchers also propose that as users pick up the pace of a rapid interaction sequence, they may learn less and make ill-considered decisions (Shneiderman, 1984).

The overwhelming majority of users prefer rapid human-computer interactions. However, the high cost of rapid response time or display rate, and the loss from increased errors must be evaluated when choosing an optimum pace (Shneiderman, 1984). One important factor that affects the desirable SRT is the short term memory capacity of users (Shneiderman, 1984; Miller, 1968). Short term memory is highly volatile, and any disruption can cause loss of memory and delays can occur when the memory needs to be refreshed. Hence, a long SRT is a disruption to the short term memory and results in a long user think time (Figure 2.1). The capacity of the short term memory depends on the task and the experience of the user among many other things, making it difficult to reach any prediction on the user's short term capacity. On the other hand, if users try to work too quickly, they may not allow sufficient time to correctly formulate a solution plan, and error rates may increase (Figure 2.2).

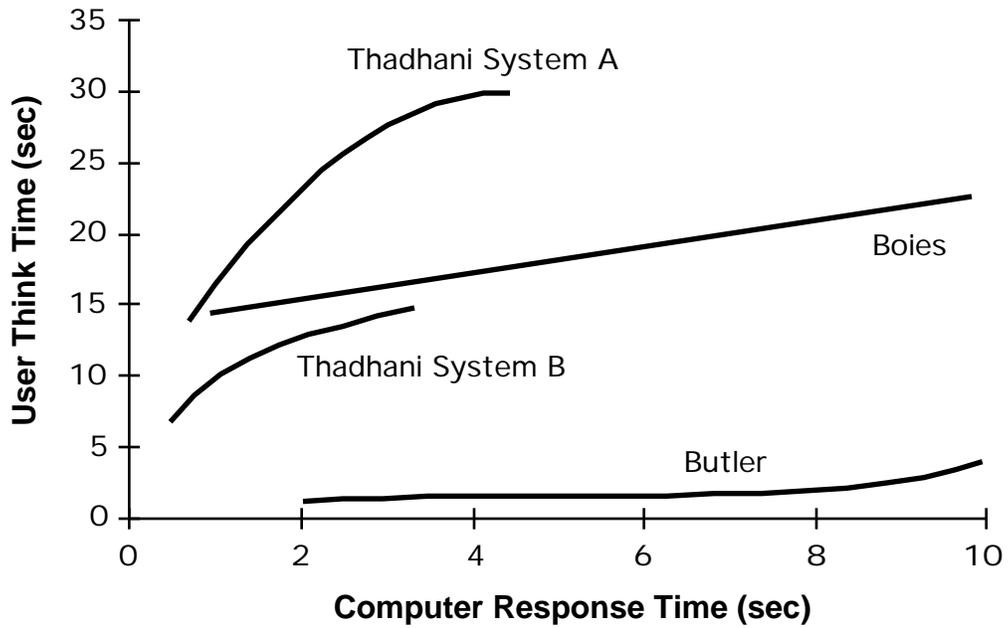


Figure 2.1 User think time as a function of computer response time in four different types of subject testings (Adapted from Shneiderman, 1992).

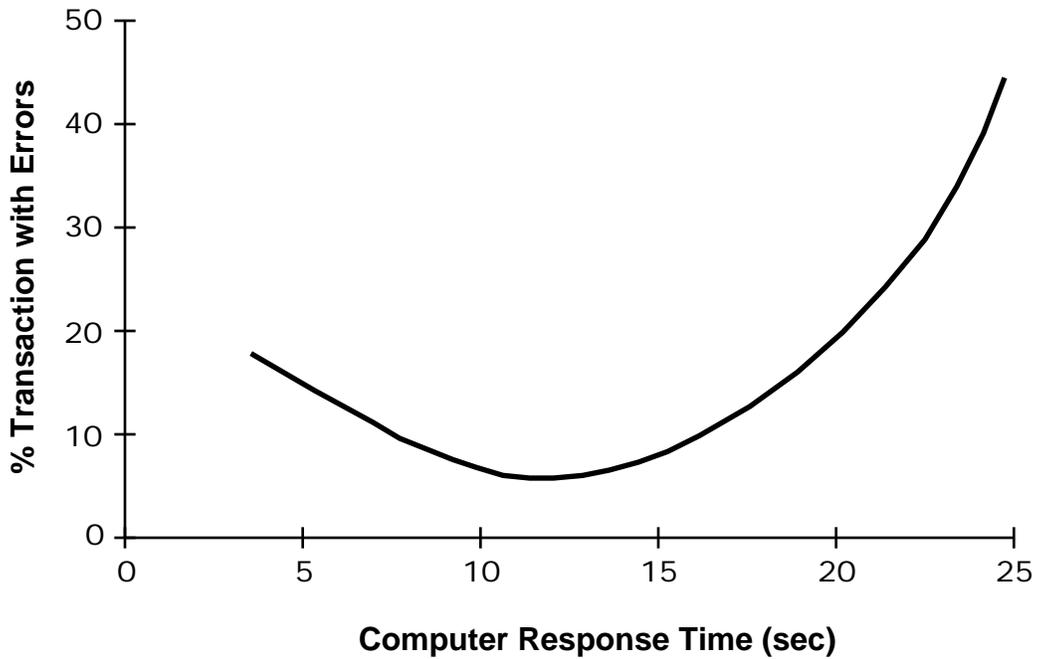


Figure 2.2 Error rates as a function of response time for a complex task (Adapted from Shneiderman, 1992).

SRT is identified as an important factor in tasks that involve problem solving. Traditionally, people believed in the “2 second” rule that states SRTs should be less than 2 seconds (Galloway, 1981). However, later studies showed that the “2 second” assumption was not always true. A fixed SRT of about 1.5 seconds had the effect of appreciably slowing down performance relative to that resulting from an SRT of less than about 1 second (Goodman and Spence 1978).

In a study by Bergman, et. al. (1981), it was found that zero second SRT led the users to commit more mistakes in completing a complex problem-solving task. The authors suggested a short SRT to complex tasks to improve user performance. Dannenbring (1984) also reached a similar conclusion after an experiment where participants engaged in “non-thinking activities”. He found that participants completed the task faster when short SRTs were introduced into the system.

In studying short-cyclic tasks, Thum, et. al. (1995) tested SRTs of 0.5, 1.5, and 4.5 seconds. The subjects’ heart rate, electrodermal responses as well as their own verbal responses were recorded. It was found that the SRT of 1.5 seconds generated least stress and was appreciated by the subjects the most. This finding demonstrated that despite the growing computing speed of today’s hardware, the notion of “the faster the better” is not always appropriate. It might be beneficial for the user if extremely short SRTs are prolonged by the software. They also found that long SRTs should be avoided.

Kuhmann (1989) conducted another short-cyclic task experiment with SRTs of 2, 4, 6, and 8 seconds. The subjects’ heart pressure, heart rate, and electrodermal activity were monitored. He found that an SRT of 6 seconds showed the lowest error rate, the shortest task completion time and lowest subject stress. On the other hand, 8 second SRT was the worst. In another study, Kuhmann, et. al. (1987) used the same experimental method except that time pressure was introduced and sometimes SRT variability existed. The time pressure seemed to enhance the efforts of subjects, causing faster average work speed and higher stress levels. The short SRTs caused the subjects to have higher blood pressure and higher headache intensity, and the long SRTs led to better performance with respect to accuracy but also caused emotional stress reactions. The researcher also found that variability in SRTs is an important factor in inducing stress.

In terms of SRT variability, Schaefer (1990) reported a study on the effects of SRT variability on human-computer interaction. SRT variability can be considered to be temporal uncertainty in work flow since the user cannot anticipate the beginning of the next work step. Schaefer found that SRT variability, especially with short SRTs, greatly affect users’ ability to predict work flow, and that the sensitivity for SRT variability decreases with increasing SRT length. The researcher recommended that software developers should design the SRT variability

so that it can be held within certain limits. Weiss, et. al. (1982) determined that variability introduces more performance errors. Martin (1973) also determined that response time variability frustrates the user.

Martin and Corl (1986) conducted SRT research on experienced subjects using Minitab. They found that at the fast end of SRT, differences in SRT in the 0.1 to 1.5 second resulted in significant productivity differences, but no significantly different productivity levels were found for SRT levels above 1.5 seconds. In addition, Martin and Corl (1986) found that decreasing SRT is beneficial for user productivity in habitual tasks, but decreasing SRT has limited utility in tasks requiring focused attentional mechanisms.

There is not just one SRT duration that is optimal for all tasks; the optimal SRT has to be determined for each type of task (Thum, et. al. 1995). Gallaway (1981) and Martin (1973) developed guidelines for the desirable SRTs for many different tasks. Although the guidelines might be somewhat obsolete, they illustrate the point that desirable SRTs may differ greatly depending on each situation. Both authors also stated that for longer response times, it is desirable for the system to give some type of feedback. They believed that appropriate feedback will compensate for longer SRTs in terms of reducing stress and increasing productivity. Martin proposed that when a system response cannot be as quick as in a human conversation, feedback should be used.

Feedback

Planas and Treurniet (1988) examined two types of user feedback provided during the system response interval. The two types of feedback were the *system feedback*, which informed users that a computer search operation was in-progress, and the *time feedback* which informed the user of the estimated time to completion of the request. The studies showed that both types of feedback produced lower ratings of user annoyance with increasing SRT than do no feedback conditions. Although the results failed to identify the better type of feedback, they concluded that continuous and rhythmic feedback made long SRTs more tolerable.

Eberts (1994) believes that the user should be given as much relevant feedback as possible or adequate for the task. Eberts further argues that for feedback to be effective, it has to be a proper one. Irrelevant feedback actually hinders performance. For an interface to be successful, the feedback has to be considered carefully. Feedback can take many forms, so the designer is faced with having to choose the right one. In his book, Eberts proposes four approaches to proper feedback design.

According to Brown (1988), when users make selections from a menu, feedback should be provided. This feedback should include both auditory and visual signals. An auditory tone or beep signals the user that an input has been accepted, and a visual feedback, such as displaying the activated area in reverse video or displaying a thick box outline around it permits the user to verify that the intended choice was activated.

Feedback should be provided to acknowledge acceptance of the input and to indicate that the system is waiting for processing to be completed (Brown 1988). A message as simple as “processing input” is acceptable in this situation although dynamic progress indicators are desirable. Dynamic progress indicators can be a numeric or graphical percent-done indicator or a counter to provide feedback on how many of a series of operations have been completed. In many situations, the preferable dynamic feedback is a count-down counter so that an estimate of completion time can be provided (Brown 1988). With proper feedback, long SRTs generate less annoyance to users.

Interface Format

There is little published work on the way in which display configuration affects problem-solving performance (Goodman and Spence, 1982). Recently, it has been widely accepted that menu selection dialogues require little or no user long-term memory to reach mastery. The menu guides the users through the operating steps and should be the display configuration of choice (Snowberry, et. al., 1983). It also is generally agreed that when the choice of topics grows and “becomes difficult to maintain under intellectual control, people form categories of similar items, creating a tree structure (Shneiderman, 1992, p. 109).” When properly designed, tree structured menu systems have the power to make large collections of data easily accessible for both novice and experienced users. The grouping of subjects can be extremely subjective and may occasionally lead to confusion, and the hierarchy structure may hinder the performance of experienced users (Shneiderman 1988). A trial and error approach with subjective testing and user feedback must be employed to improve the success rate of the tree structure (Shneiderman, 1992). Besides suggesting that the design should be iterative, Lee, et. al. (1984) also noticed that novice users can judge the quality of the tree structure menu design better than experienced users.

The two main features of the tree structure are the depth, or number of levels, and breadth, or number of items per level. If the tree structure has more breadth, then fewer levels are required for a given choice. Several studies made on depth-breadth trade-off showed evidence that breadth should be preferred over depth (Kiger, 1984; Landauer and Nachbar, 1985; Snowberry, et. al., 1983). Kiger (1984) actually determined that the depth should be limited to three levels and that when the depth goes over this number, there is a greater chance of users becoming disoriented. Landauer and Nachbar (1985) also found that there was a clear significant

cost associated with greater depth. Snowberry, et. al. (1983) found that as depth decreased, search speed and accuracy improved. The navigation problem is believed to become more treacherous as the depth of the hierarchy increases causing increased error rates (Paap and Roske-Hofstrand, 1986). If a menu contains many levels, the menu is said to be funneling. However, if the menu has excessive breadth, crowding occurs, and might fill up all the available space on a panel making it extremely difficult to find the desired item (Paap and Roske-Hofstrand, 1986). Lee and MacGregor (1985) determined that for systems with short SRTs, funneling generates fast overall processing time. When the research was further extended by Paap and Roske-Hofstrand in 1986, they determined that as users gain experience, the breadth and depth trade-off on processing time becomes insignificant.

In experimenting with some alternative designs Shneiderman (1992) recommended great breadth at the root and at the leaves. This makes it possible to decrease the number of levels, increasing user performance. When Shneiderman (1992) introduced stress to users to make them work quickly, he found that the users committed the least number of errors on systems with greater breadth.

Shneiderman (1992) proposed a four point guideline in menu tree design. The guideline includes (1) Create groups of logically similar items. (2) Form groups that cover all possibilities. (3) Make sure that items are nonoverlapping, and (4) Use familiar terminology, but ensure that items are distinct from one another. Other details to watch for include consistent formats for predictability, added easy instructions, error messages, and status report or intermediate feedback.

As the depth of the tree grows, Shneiderman (1992) suggests the incorporation of a menu map or a printed index of terms that is easier to scan than a series of screen displays.

CHAPTER 3. EXPERIMENTAL METHOD

Facilities and Apparatus

The experiment was conducted in the Displays and Controls Laboratory (DCL) at Virginia Polytechnic Institute and State University. One of the rooms in the DCL was set up to have the atmosphere of a living room to simulate the anticipated real life conditions (Figure 3.1). A multimedia software package (Macromedia, Director 5.0) was used to prepare a prototype for the system. This prototype ran on a computer (Apple, Power Macintosh 8500/120) that was linked to a color television (Sony, KV-32S10) through a digital I/O board (National Instruments, PCI-DIO-96). The participants used a programmable Universal Remote Control (Tele-Art, RC-100) (Figure 3.2) to browse the WWW (prototype) on the television. This remote control sent signals to an infrared receiver (Velleman-kit, 15-channel IR kit) that was connected to the digital I/O board in the computer. Figure 3.3 illustrates the setup for the experiment.

Participants

Sixteen individuals participated in this study. They were prescreened for experience with computers and browsing the WWW (Appendix A). Eight of the participants were experienced, while the remaining eight were inexperienced. Inexperienced participants' computer knowledge was limited to occasional use of word processors and they had never or rarely been on the WWW. Experienced participants used computers extensively (e.g., word processing, spreadsheets, programming) had certain WWW sites that they visited regularly, and were on the WWW almost every day. In addition, each participant was tested for far acuity to ensure that they could see the television screen clearly.

Experimental Design

The experiment was a four-factor, mixed-model design, as illustrated in Figure 3.4. Participants were nested within Experience Level (Exp) and the three within-factors considered were: Interface Format (I), Feedback Style (F), and System Lag (L). There were a total of 30 treatment combinations; 15 treatment combinations for each interface format. The participants were presented with all the treatment combinations in one interface format before they were presented with the treatment combinations in the other one. The presentation order of the different interface formats were counterbalanced in order to prevent any carry over effects. The treatment combinations within each interface format were assigned randomly to each participant.

Each participant was asked to complete a total of 42 tasks, of which 12 were for the practice session (Appendix B). The tasks consisted of finding Web pages in the prototype by navigating through the tree structure.

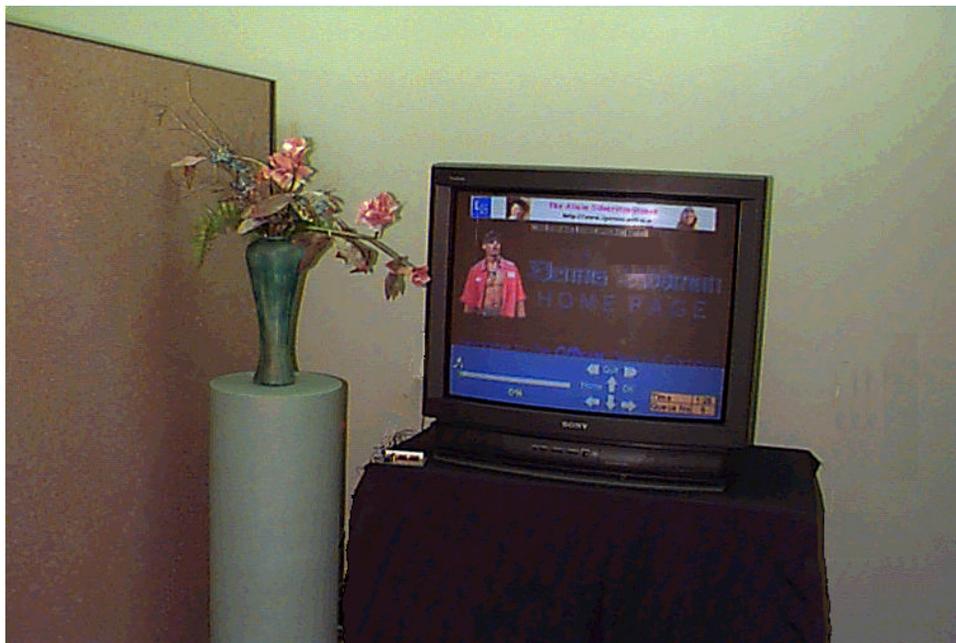


Figure 3.1 Setup of the experiment room.



Figure 3.2 The universal remote control used in the experiment.

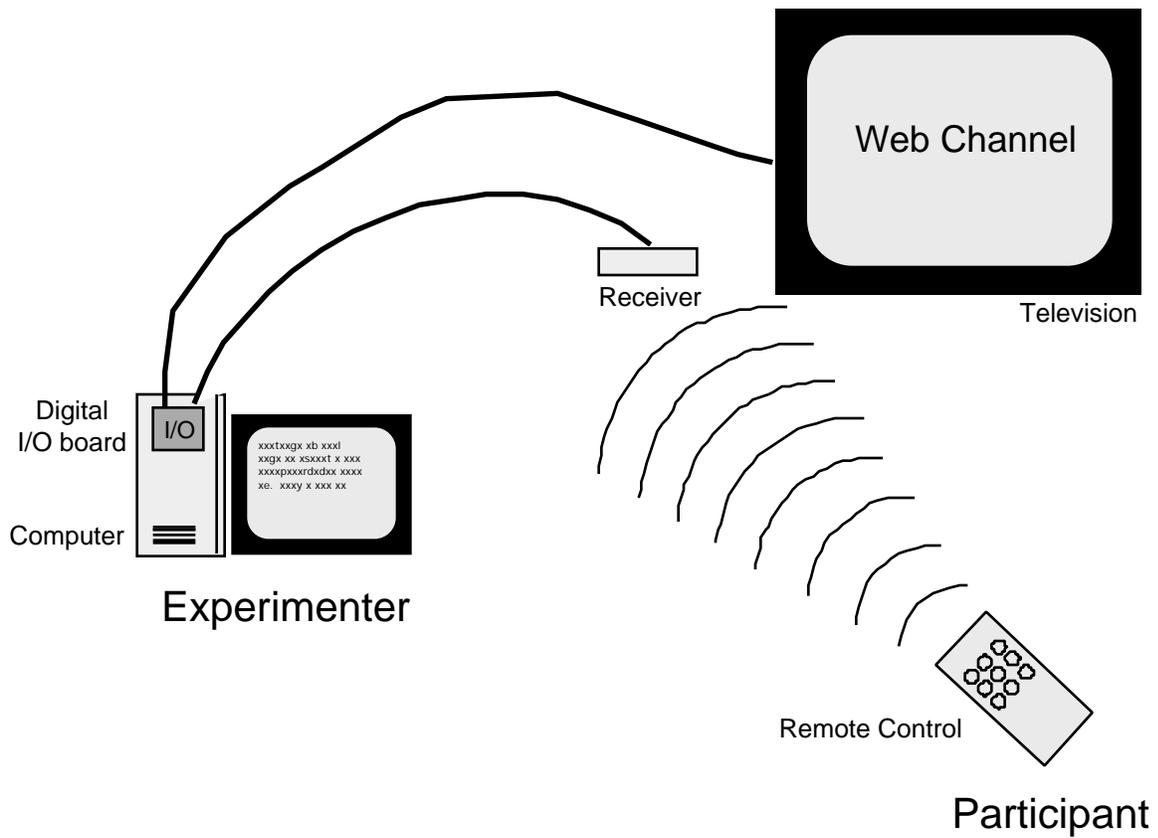


Figure 3.3 Setup for the experiment.

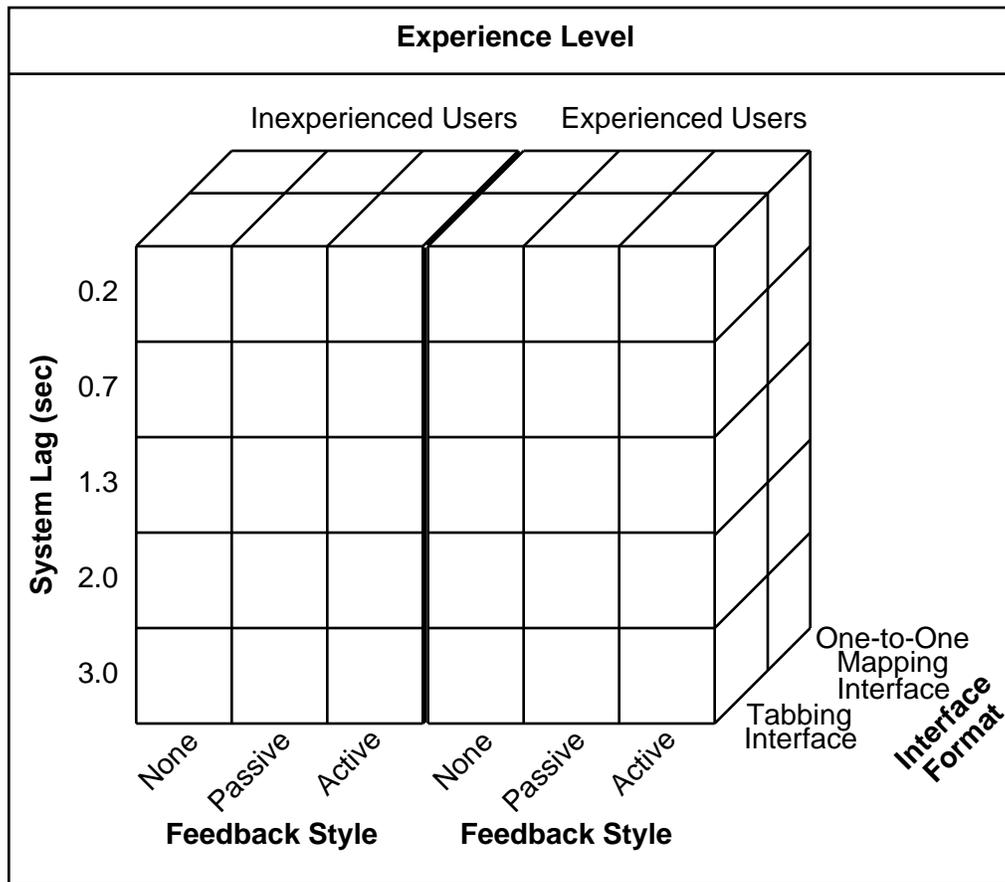
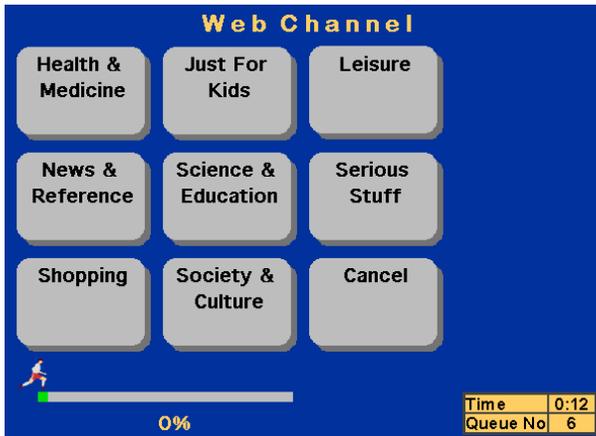


Figure 3.4 Experimental design.

Independent Variables

Interface Format

There were two interface stages in the WWW browser system. The first one was the indexing stage that lasted until a user reached a Web page. The second stage was when the user was browsing the WWW by choosing certain links in a Web page (Figure 3.5c). This study was restricted to the first interface stage (indexing stage). There were two variations of the interface format for the design of the indexing stage: the one-to-one mapping interface format (Figure 3.5a) and the tabbing interface format (Figure 3.5b). Both of these interface formats were tree menu structures. The participants navigated through these indexed tree structures and were guided to web sites.



(a) Indexing stage with the One-to-One Mapping Interface.



(b) Indexing stage with the Tabbing Interface.



(c) Web stage.

Figure 3.5 The two interface stages.

Both interface formats had some common attributes. The time/queue box located at the bottom right corner indicated how long a user had been using the system and how many more people somebody needed to wait for before the service was available to him/her. As users navigated in the tree structure, the path they followed appeared on the right side of the screen informing the users about how they reached that particular topic.

One-to-one mapping interface had nine choices on the screen arranged in a 3x3 grid corresponding to the layout of the number buttons one (1) through nine (9) on the remote control. Figure 3.6a illustrates how the remote control worked with the one-to-one mapping interface. There were up to nine categories per level in the tree structure for the one-to-one mapping interface and the tree structure was up to six levels deep. Figure 3.7 shows the step by step process of reaching a Web site using the one-to-one mapping interface. One of the nine buttons on the screen was a dedicated *cancel* button, leaving up to eight categories per level. The *cancel* button led to another screen that made it possible to go to the beginning of the index, back and forward in the indexing structure, and quit the service (Figure 3.8).

The other interface format, tabbing interface, had up to 18 categories per level and the tree structure was up to four levels deep. Each one of the nine buttons on the remote control had specific functions (Figure 3.9a) and the participants used the nine buttons to tab through the choices on the screen. Figure 3.6b illustrates how the remote control worked with the tabbing interface. Figure 3.10 shows the step by step process of reaching a Web site using the tabbing interface.

Once the participants reached a desired topic, a list of web pages were displayed on the screen with a brief description of each web site below the horizontal line (Figure 3.7d, Figure 3.10c). From this point on, the two interface formats were the same: the nine buttons had the same functions as in the tabbing interface (Figure 3.9a). The indexing structure ended when the participants chose a web page from the list. Although it was not used in this experiment, Figure 3.9b shows the functions of the nine buttons during the Web stage.

Since the participants were asked to complete 42 tasks during the experiment, there were only 42 Web sites in the prototype. This made it possible for participants to come to a *dead end* in the index. In such a case they were presented with the screen in Figure 3.11a. A similar situation occurred if the participants pressed *quit*; that is, they were presented with the display in Figure 3.11b.

Feedback Style

Each time participants entered a command associated with a screen change (e.g., changing a level in the tree structure, going to the *cancel* screen, selecting a Web page), feedback was provided indicating that the system was processing the request. The reason for feedback was to alleviate the user stress associated with network delays for downloading new pages. During the indexing stage, each button press led to a screen change in the one-to-one mapping interface. Therefore, each button press was followed by network download time. In the tabbing interface; however, only a few of the button presses were followed by network download time since most of the button presses were for tabbing through the topic choices on the screen. In this study, network download time was set to five seconds; hence feedback was displayed for five seconds before a screen change could occur.

There were three types of Feedback Styles (F) in the study. The first style did not provide the participants with any feedback (Figure 3.12a). In the second style (passive feedback), the feedback was an icon of a man running in place, which did not present any indication of estimated time to complete a request (Figure 3.12b). The last type of feedback (active feedback) informed the participants about how much of their request had been processed by displaying the running man icon along a bar and the percentage of request completed (Figure 3.12c).

System Lag

System lag is defined as the time between an input from a user and the response from the system. There were five levels of system lag in the experiment ranging from 0.2 seconds to 3.0 seconds: 0.2, 0.7, 1.3, 2.0, and 3.0.

Dependent Variables

Task completion time was measured in seconds and the number of errors committed was recorded for each task. An error was recorded each time participants took a wrong path in the tree structure. If they were able to reverse the wrong path, the action was counted as one error. However, if the participants decided to pursue another incorrect path on the way back from the wrong path, the action was considered as a second error. If participants were not able to complete a task in five minutes, they were asked to skip that task and start the next one. At the end of each task, the participants were asked to complete rating scales to evaluate the responsiveness of the system, difficulty of the task, acceptability of the feedback style and the overall usability of the system. Furthermore, they were asked to complete Likert-type scales

about each interface format. At the end of the experiment, participants were asked to mark their preference of feedback style and interface format. Refer to Appendix C for the subjective evaluation questions.

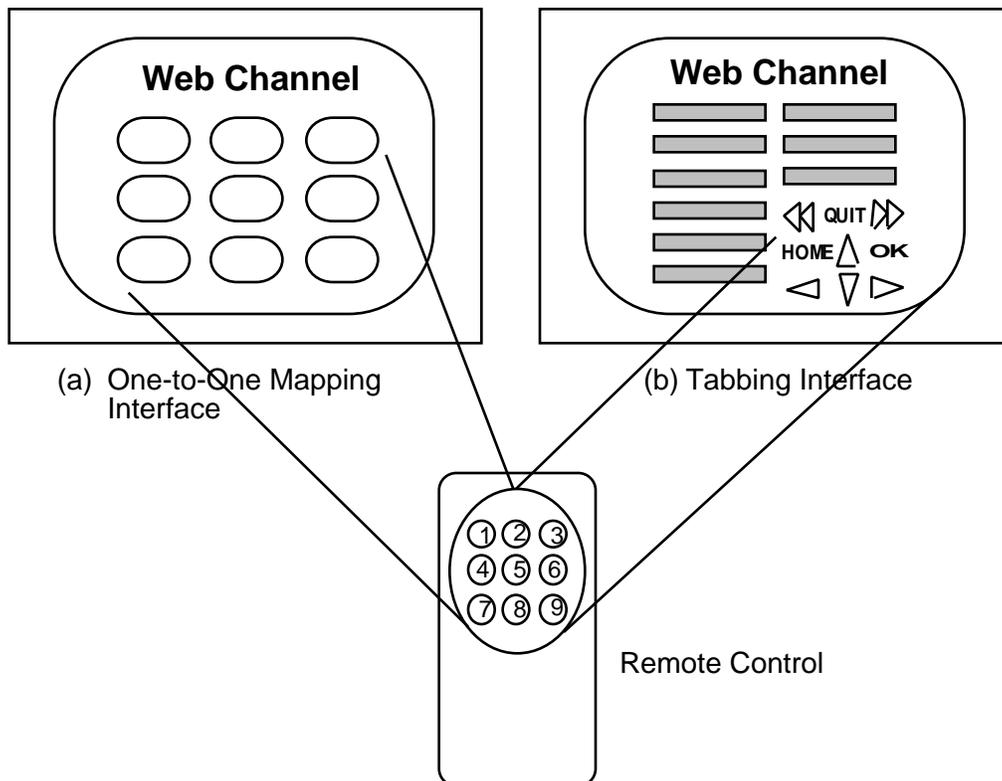
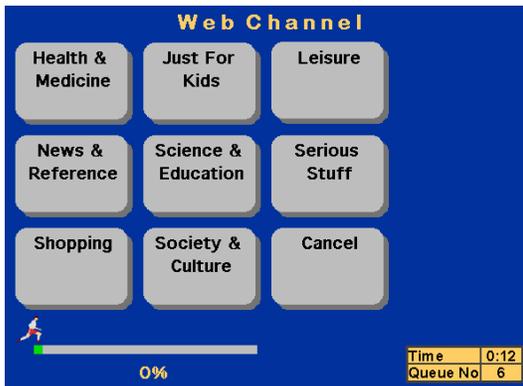
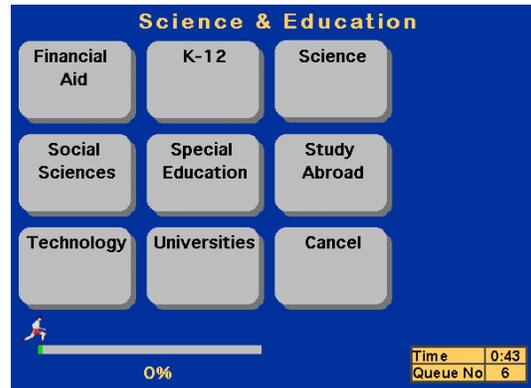


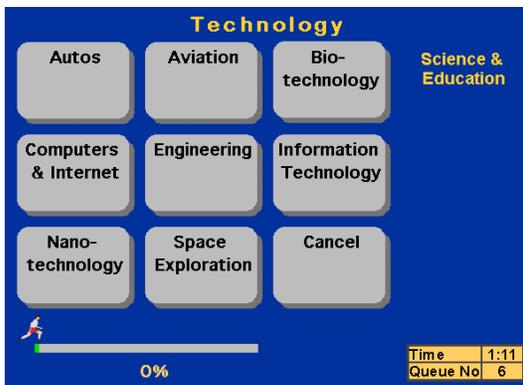
Figure 3.6 Illustration of how the nine buttons on the remote control work with the two interface formats.



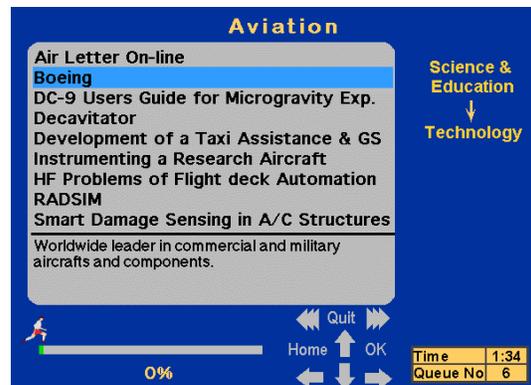
(a)



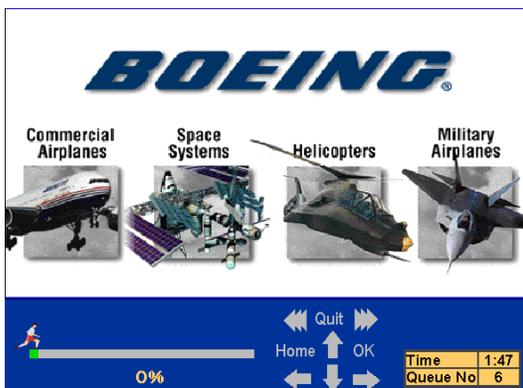
(b)



(c)



(d)



(e)

Figure 3.7 Step-by-step procedure for reaching the Boeing web site using the one-to-one mapping interface format.

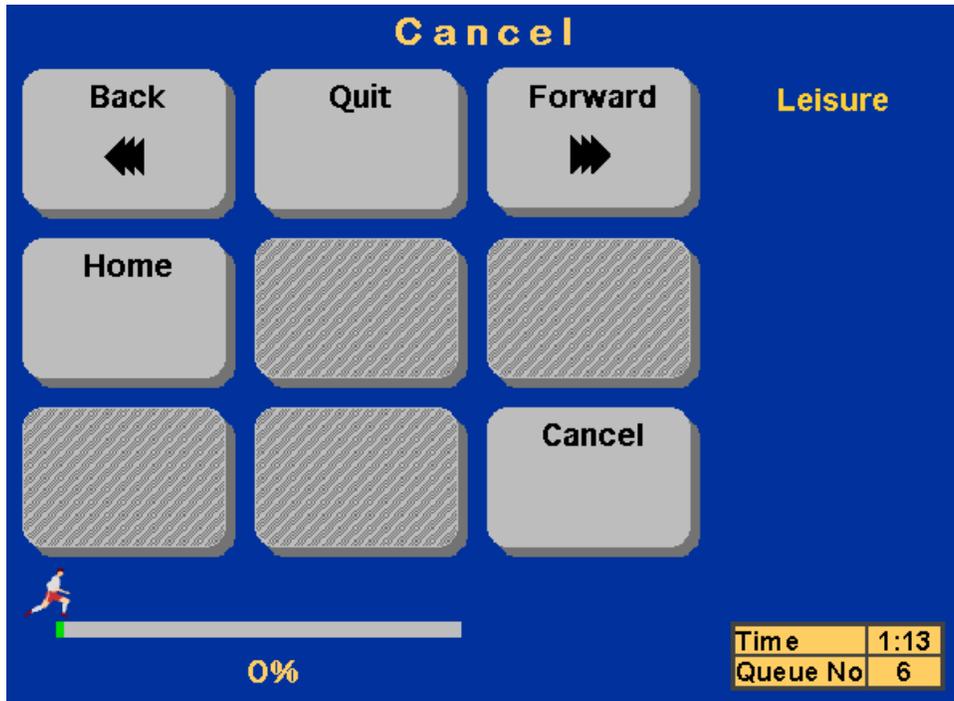


Figure 3.8 Cancel screen in the one-to-one mapping interface.

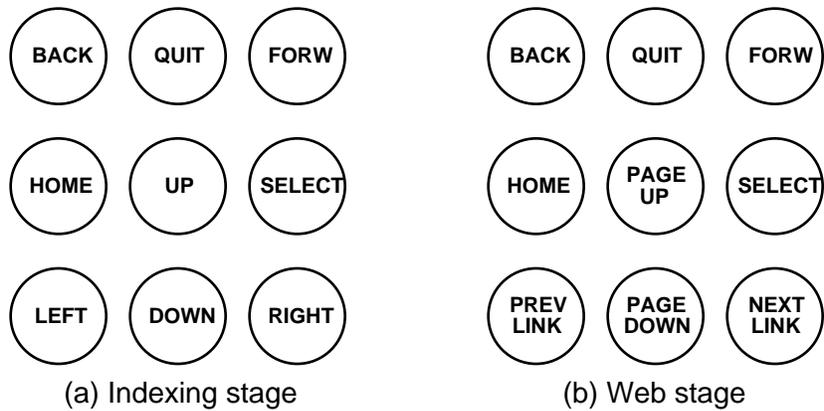
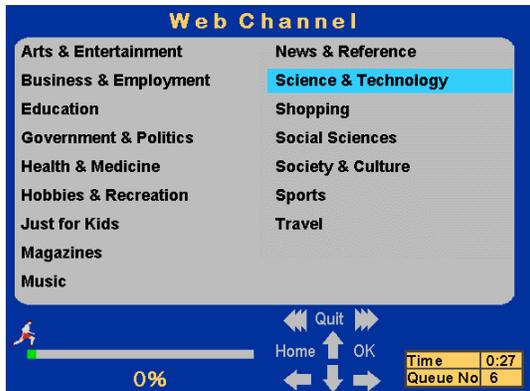
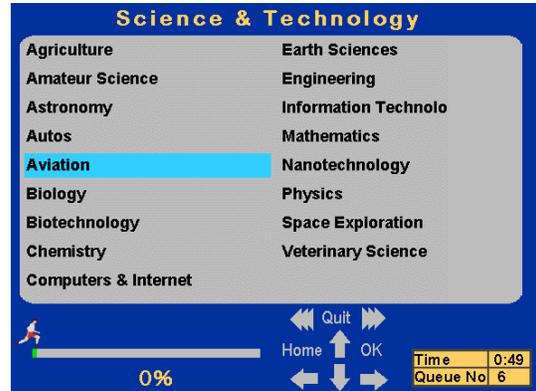


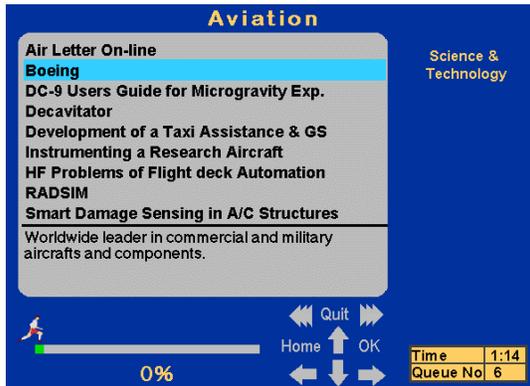
Figure 3.9 Functions of the buttons on the remote control.



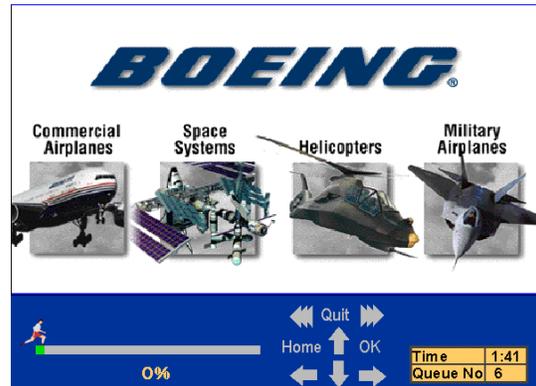
(a)



(b)

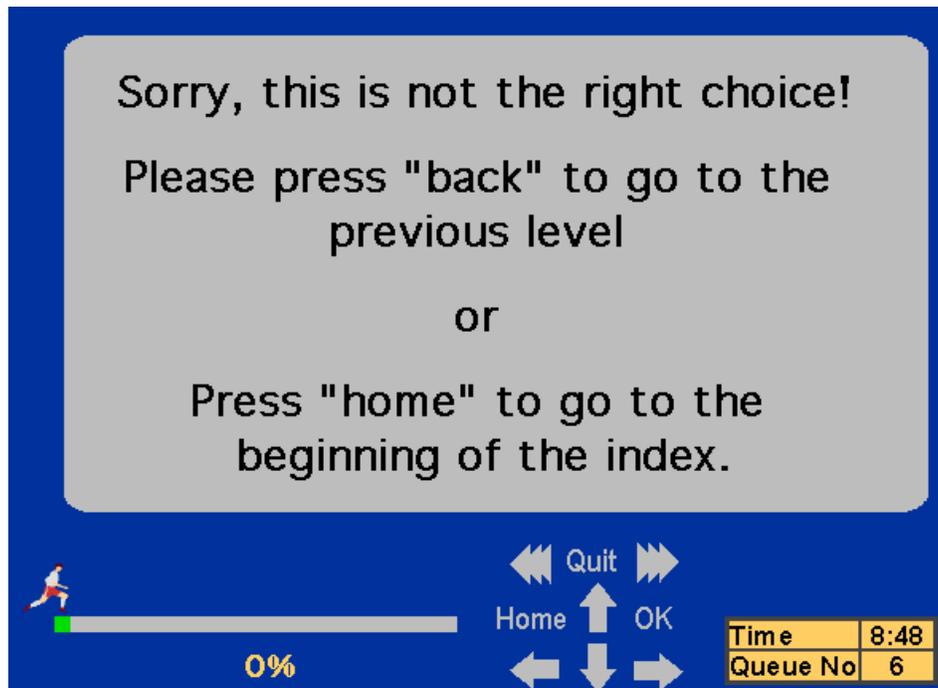


(c)

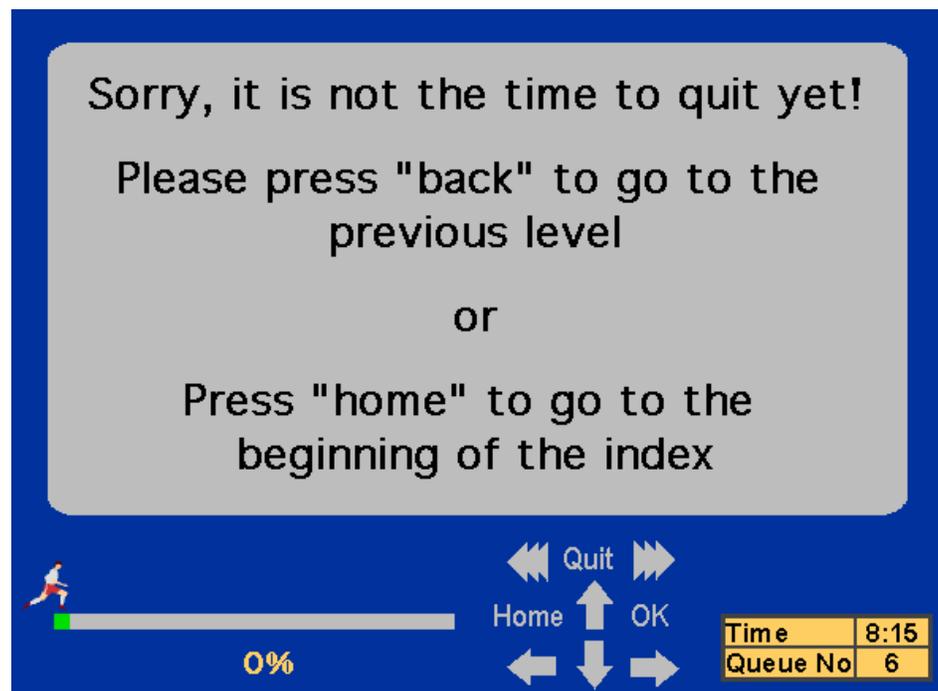


(d)

Figure 3.10 Step-by-step procedure for reaching the Boeing web site using the tabbing interface format.



(a) Dead end screen.



(b) Quit screen.

Figure 3.11 Screens that appeared (a) when participants chose a topic that did not lead to web pages and (b) when participants pressed “quit”.



(a) No feedback.



(b) Passive feedback.



(c) Active feedback.

Figure 3.12 Three different types of feedback styles.

Mathematical Model

Figure 3.13 illustrates the timeline for the two interface formats. Assuming user response time and user think time are negligible, and tasks are completed without errors, average task completion time depends on the system lag and the network download time. Based on the timelines in Figure 3.13, average task completion time can be expressed as follows for the one-to-one mapping interface:

$$T = L_S a_t + L_N a_t + L_S a_w + L_S b_w + L_N b_w \quad (\text{Eq.1})$$

$$T = L_S (a_t + a_w + b_w) + L_N (a_t + b_w) \quad (\text{Eq.2})$$

where T is average task completion time;

L_S is system lag time;

L_N is network download time;

a_t is average number of button presses in tree structure;

a_w is average number of button presses in Web page list, and

b_w is average number of OK presses in Web page list.

Substituting the values from the 30 tasks completed in the study, equation 2 becomes:

$$T = L_S (3.77 + 3.83 + 1) + L_N (3.77 + 1) \quad (\text{Eq.3})$$

$$T = 8.6L_S + 4.77L_N \quad (\text{Eq.4})$$

Average task completion time for the tabbing interface can be expressed with the following equation:

$$T = L_S a + L_N b \quad (\text{Eq.5})$$

where a is average number of button presses, and

b is average number of OK presses.

Substituting the values from the 30 tasks completed in the study, equation 5 becomes :

$$T = 16.17L_S + 3.47L_N \quad (\text{Eq.6})$$

Equations 4 and 6 reveal that if network download time is constant, increasing the system lag causes the task completion time for the tabbing interface to increase by a factor of 1.88 compared to the one-to-one mapping interface. However, if system lag is constant, it can be seen that longer network download times increase task completion time for the one-to-one mapping interface by a factor of 1.375 compared to the tabbing interface. In this study, network download time was set to five seconds and system lag time had five different values between 0.2 and three seconds. Figure 3.14a reveals the implications of equations 4 and 6 when network download time is equal to five seconds. Figure 3.14b illustrates task completion time as a function of network download time when system lag time is 0.2 seconds. Since network download time is highly variable in the Internet, these equations can be used as a guide for developing user interface designs for systems that enable access to the Internet.

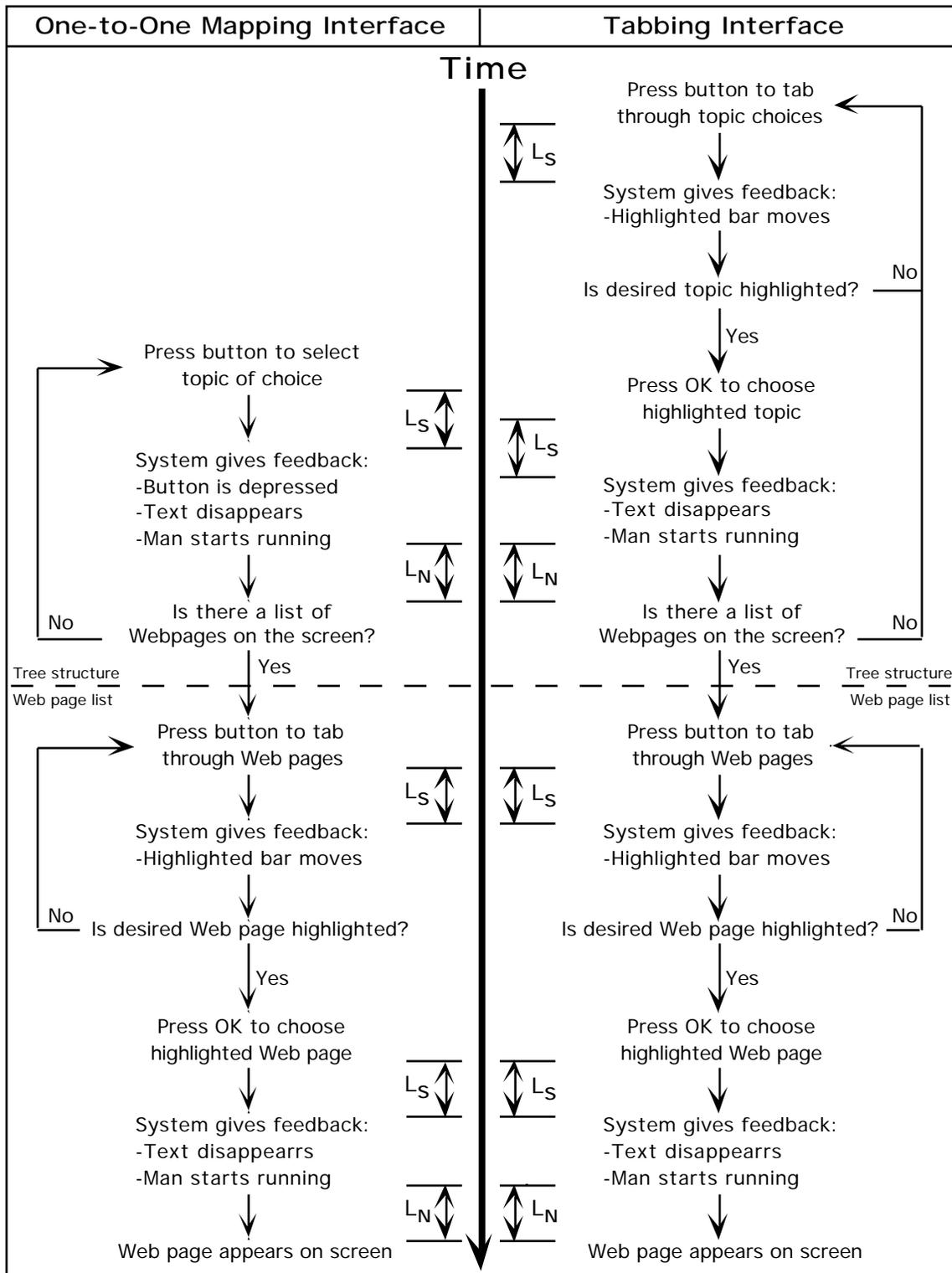
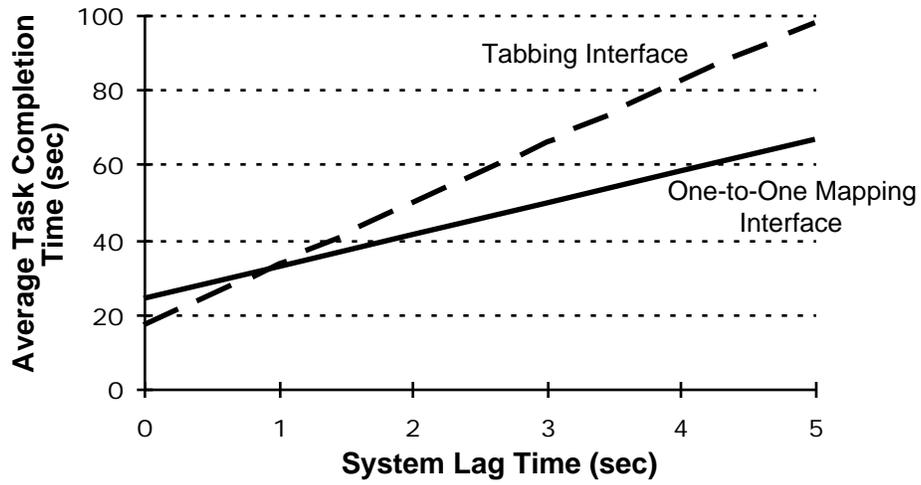
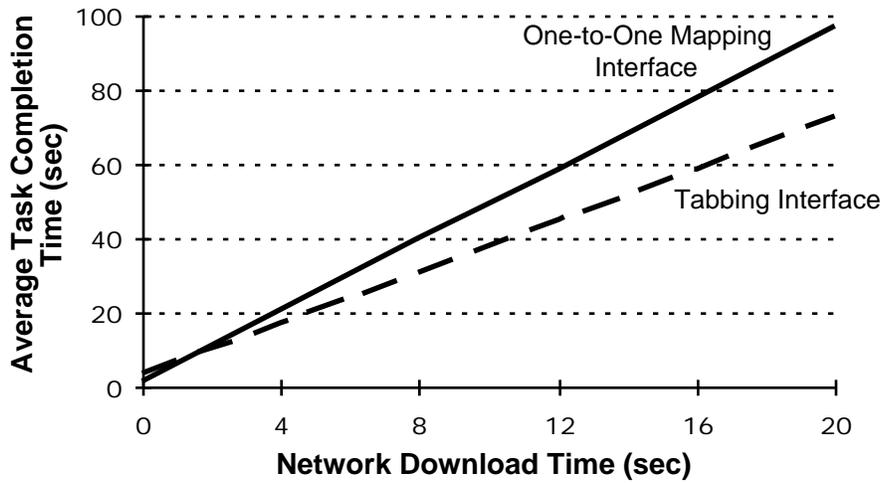


Figure 3.13 Timeline for the two interface formats. L_S is system lag time and L_N is network download time (not drawn to scale).



(a) Network download time is five seconds.



(b) System lag time is 0.2 seconds.

Figure 3.14 Average task completion time as a function of (a) System lag time and (b) Network download time.

Experimental Procedure

Pre-experimental Training

Participants were trained on the system before data collection started. After the training, they were asked to complete 12 tasks similar to the ones used in the actual data collection stage (Appendix B). They were exposed to all the different types of interface format, feedback style, and system lag. The purpose of the pre-experimental training was to familiarize them with the system to prevent any learning effects during data collection.

Experimental Session

Participants were asked to complete 30 tasks during the data collection stage of the experiment (Appendix B). These tasks originated from the first screen of the indexing stage and ended when the participants reached a Web site. There was a set time limit of five minutes for task completion. If the participants could not complete a task in five minutes, they were asked to proceed to the next task. After performing each task, the participants completed rating scales to provide their subjective opinion on the responsiveness of the system, difficulty of the task, acceptability of the feedback style, and the usability of the system (Appendix C). The participants completed 15 tasks with one interface format before proceeding to the next set of 15 tasks with the other interface format. The tasks and the corresponding feedback styles and system lags within each interface format were randomly assigned to each participant. Participants evaluated each interface format by completing Likert-type rating scales (Appendix C). At the end of the experiment, participants indicated their preference of feedback style and interface format (Appendix C).

Data Analysis

Analysis of Variance (ANOVA) techniques were used to analyze the task completion time, number of errors, and the subjective evaluation ratings. The Greenhouse-Geisser correction was applied to all significant main effects and interactions. Significant effects were explored through Newman-Keuls post-hoc analysis. Furthermore, subjective rating evaluations on each interface format and participants' preferences on feedback style and interface format were analyzed using descriptive statistics.

CHAPTER 4: RESULTS

Task Completion Time

As demonstrated by the ANOVA summary table (Table 4.1), the main effects of interface format and system lag had significant impact on task completion time. In terms of interface format, there was an increase of 14% in task completion time with the tabbing interface compared to the one-to-one mapping interface (Figure 4.1). As presented in Figure 4.2, except for the system lag of 1.3 seconds, task completion time increased as system lag increased. The Newman-Keuls post-hoc analysis (Table 4.2) indicates that 0.2 second and 0.7 second system lags resulted in significantly shorter task completion times than did the system lags of 1.3 and 3.0 seconds.

Number of Errors

The ANOVA summary table, presented in Table 4.3, reveals that the experience level of the participants and the interface format had significant effect on the number of errors committed. On average, the experienced users committed 33% more errors than did the inexperienced users (Figure 4.3). As shown in Figure 4.4, there was a 53% reduction in errors when participants used the tabbing interface.

Subjective Ratings

System Lag Time

The ANOVA summary table for the System Lag subjective ratings is presented in Table 4.4. The main effects of Interface Format and System Lag were found to be significant, as well as the interaction between Interface Format and Feedback Style.

Participants were more tolerant to system lag with the one-to-one mapping interface than with the tabbing interface. The ratings for system lag were 22% higher for the one-to-one mapping interface than they were for the tabbing interface (Figure 4.5). As shown in Figure 4.6, 0.2 seconds of system lag received the highest ranking and 3.0 seconds the lowest. Table 4.5 illustrates the results of the Newman-Keuls post-hoc analysis that indicate participants' ratings increased significantly as system lag became shorter.

Figure 4.7 shows the average System Lag rating for the Interface Format \times Feedback Style interaction. One-to-one mapping interface with the active feedback received the highest rating while the tabbing interface with no feedback received the lowest rating. However, the Newman-Keuls post-hoc analysis (Table 4.6) revealed that system lag ratings were insignificant for this interaction.

Table 4.1 ANOVA Summary Table for Average Task Completion Time.

Source	df	SS	MS	F	p	p _{G-G}
<u>Between-Subjects</u>						
Experience (Exp)	1	15675	15675	2.203	0.1599	
Subjects/Experience (S/Exp)	14	99606	7115			
<u>Within-Subjects</u>						
Interface Format (I)	1	31639	31639	5.987	0.0282*	<0.10
I \times Exp	1	2435	2435	0.3461	0.5082	
I \times S/Exp	14	73997	5285			
Feedback Style (F)	2	22440	11220	1.470	0.2471	
F \times Exp	2	5850	2925	0.383	0.6853	
F \times S/Exp	28	213691	7632			
System Lag (L)	4	141829	35457	4.406	0.0036*	<0.10
L \times Exp	4	4049	1012	0.126	0.9725	
L \times S/Exp	56	450645	8047			
I \times F	2	4594	2297	0.466	0.6323	
I \times F \times Exp	2	31420	15710	3.186	0.0567	>0.10
I \times F \times S/Exp	28	138067	4931			
I \times L	4	22296	5574	1.000	0.4153	
I \times L \times Exp	4	15936	3984	0.714	0.5859	
I \times L \times S/Exp	56	312269	5576			
F \times L	8	83386	10423	1.750	0.0946	>0.10
F \times L \times Exp	8	108730	13591	2.282	0.0266	>0.10
F \times L \times S/Exp	112	667021	5956			
I \times F \times L	8	16557	2070	0.309	0.9613	
I \times F \times L \times Exp	8	14988	1874	0.279	0.9716	
I \times F \times L \times S/Exp	112	751259	6708			
<u>Total</u>	479	3228377				

*Indicates effects determined to be statistically significant after applying the Greenhouse-Geisser correction (p=0.10).

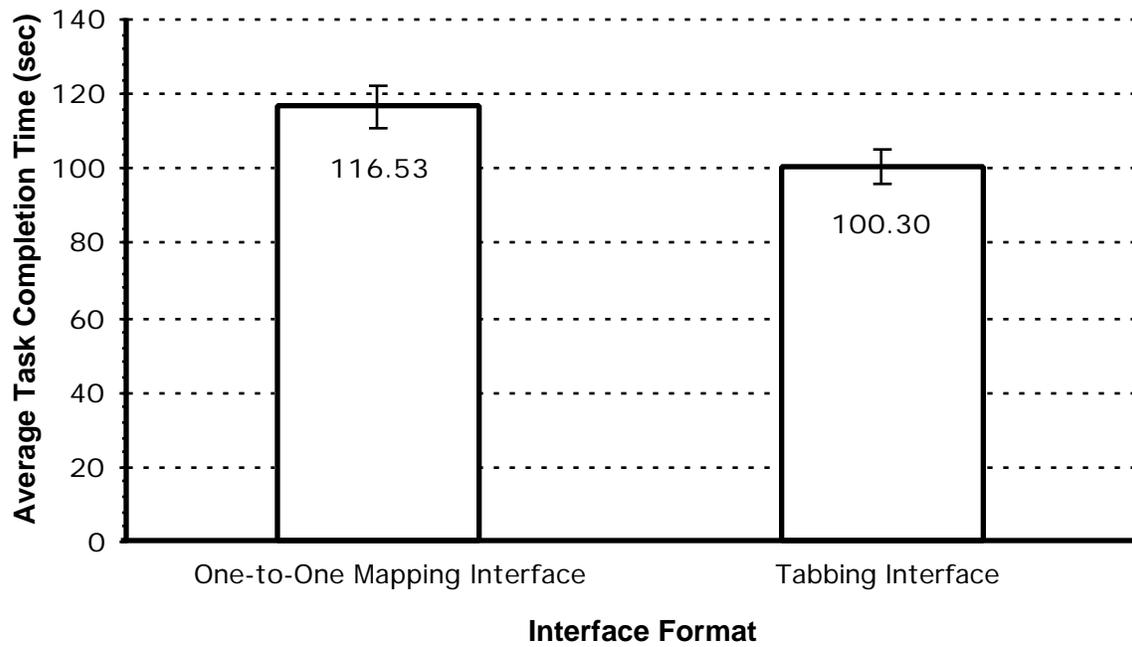


Figure 4.1 Effect of Interface Format on task completion time. Error bars indicate ± 1 standard error of the mean (N=240).

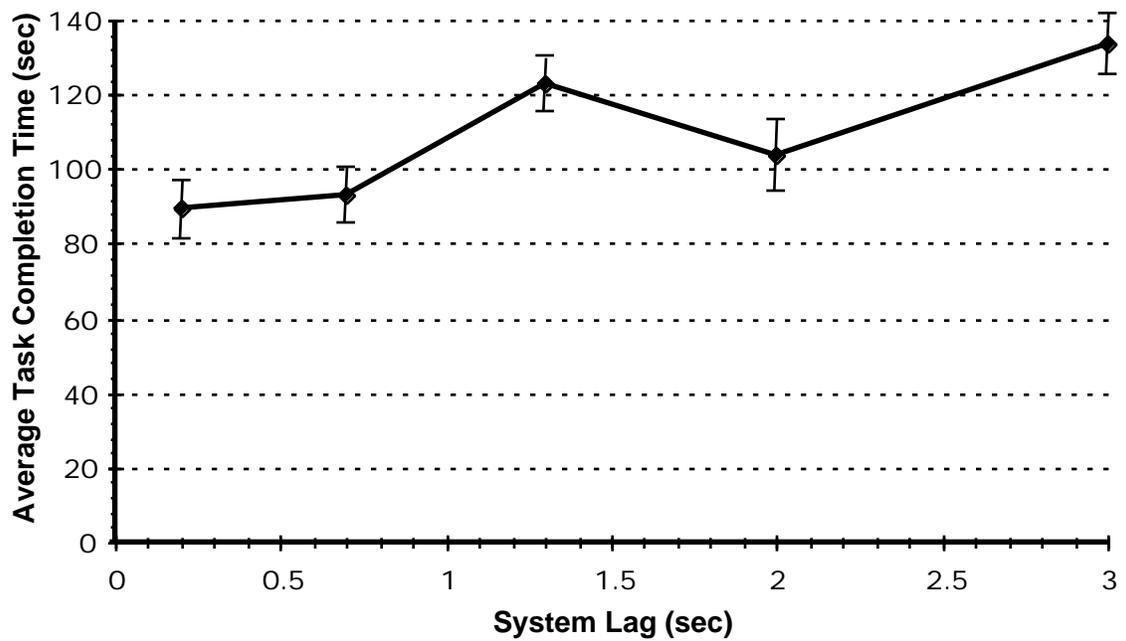


Figure 4.2 Effect of System Lag on task completion time. Error bars indicate ± 1 standard error of the mean (N=96).

Table 4.2 Newman-Keuls Comparison of Task Completion Time among System Lag Times.

System Lag (sec)	Mean	Group
3.0	134.0	A
1.3	123.0	A B
2.0	104.0	B C
0.7	98.2	C
0.2	89.3	C

Groups with the same letter are not significantly different at $p=0.10$

Table 4.3 ANOVA Summary Table for Average Number of Errors.

Source	df	SS	MS	F	p	p_{G-G}
<u>Between-Subjects</u>						
Experience (Exp)	1	14.700	14.700	5.444	0.0351*	<0.10
Subjects/Experience (S/Exp)	14	37.800	2.700			
<u>Within-Subjects</u>						
Interface Format (I)	1	49.408	49.408	13.325	0.0026*	<0.10
I×Exp	1	8.008	8.008	2.160	0.1638	
I×S/Exp	14	51.917	3.708			
Feedback Style (F)	2	11.717	5.858	1.390	0.2657	
F×Exp	2	3.800	1.900	0.451	0.6415	
F×S/Exp	28	117.950	4.213			
System Lag (L)	4	18.992	4.748	1.203	0.3198	
L×Exp	4	1.175	0.294	0.074	0.9898	
L×S/Exp	56	221.033	3.947			
I×F	2	2.817	1.408	0.484	0.6214	
I×F×Exp	2	15.267	7.633	2.626	0.0901	>0.10
I×F×S/Exp	28	81.383	2.907			
I×L	4	14.133	3.533	1.114	0.3591	
I×L×Exp	4	9.617	2.404	0.758	0.5570	
I×L×S/Exp	56	177.583	3.171			
F×L	8	41.533	5.192	1.774	0.0895	>0.10
F×L×Exp	8	48.950	6.119	2.091	0.0424	>0.10
F×L×S/Exp	112	327.717	2.926			
I×F×L	8	10.017	1.252	0.337	0.9499	
I×F×L×Exp	8	5.233	0.654	0.176	0.9937	
I×F×L×S/Exp	112	415.617	3.711			
<u>Total</u>	479	1686.367				

*Indicates effects determined to be statistically significant after applying the Greenhouse-Geisser correction ($p=0.10$).

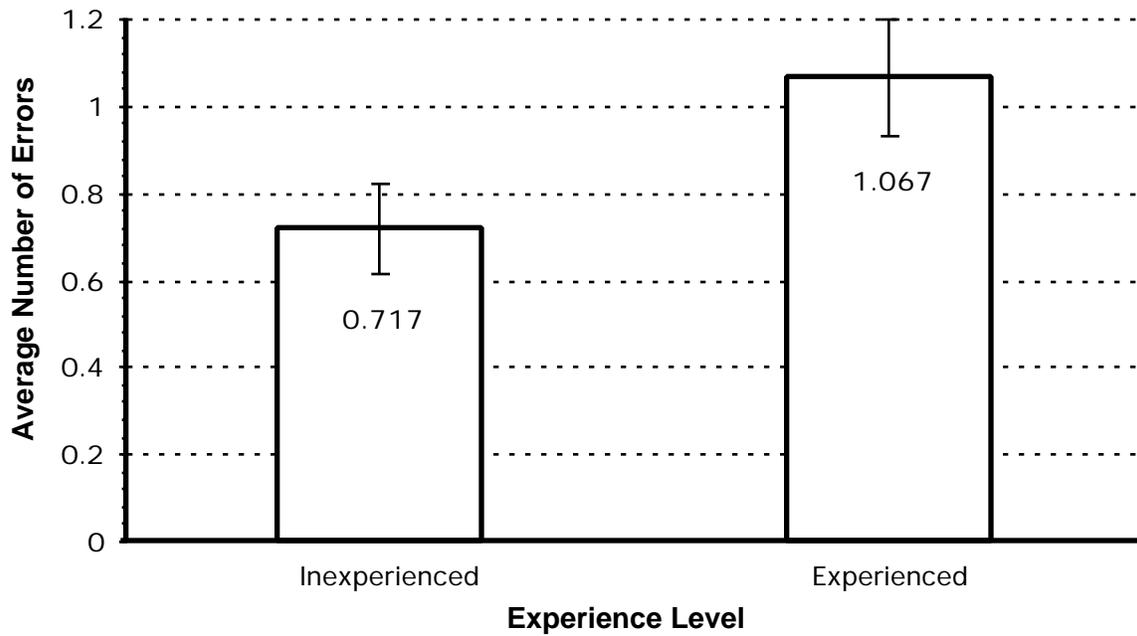


Figure 4.3 Effect of Experience Level on number of errors committed. Error bars indicate ± 1 standard error of the mean (N=240).

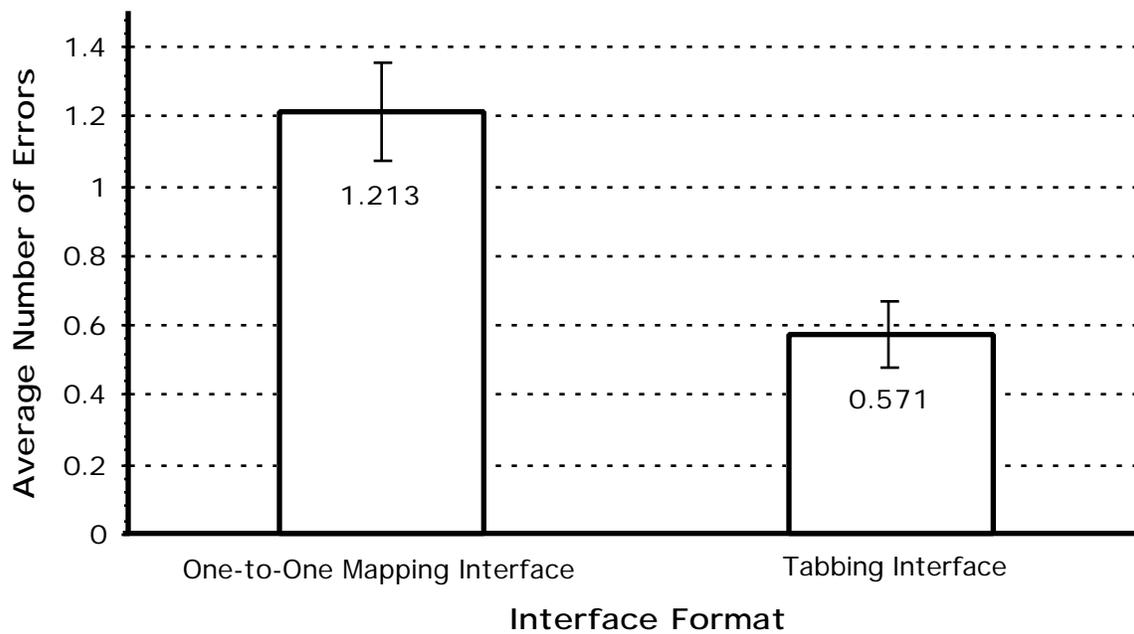


Figure 4.4 Effect of Interface Format on number of errors committed. Error bars indicate ± 1 standard error of the mean (N=240).

Table 4.4 ANOVA Summary Table for Subjective Ratings on System Lag.

Source	df	SS	MS	F	p	p_{G-G}
<u>Between-Subjects</u>						
Experience (Exp)	1	4.219	4.219	0.158	0.6970	
Subjects/Experience (S/Exp)	14	374.496	26.750			
<u>Within-Subjects</u>						
Interface Format (I)	1	31.519	31.519	27.242	0.0001*	<0.10
I×Exp	1	1.519	1.519	1.313	0.2711	
I×S/Exp	14	16.196	1.157			
Feedback Style (F)	2	1.462	0.731	1.684	0.2039	
F×Exp	2	0.262	0.131	0.302	0.7417	
F×S/Exp	28	12.142	0.434			
System Lag (L)	4	141.283	35.321	20.183	0.0000*	<0.10
L×Exp	4	14.292	3.573	2.042	0.1009	>0.10
L×S/Exp	56	98.025	1.750			
I×F	2	307.163	153.581	72.650	0.0000*	<0.10
I×F×Exp	2	11.512	5.756	2.723	0.0831	>0.10
I×F×S/Exp	28	59.192	2.114			
I×L	4	3.492	0.873	0.822	0.5167	
I×L×Exp	4	3.950	0.988	0.930	0.4532	
I×L×S/Exp	56	59.492	2.114			
F×L	8	12.829	1.604	2.030	0.0491	>0.10
F×L×Exp	8	5.446	0.681	0.862	0.5508	
F×L×S/Exp	112	88.525	0.790			
I×F×L	8	8.046	1.006	0.514	0.8438	
I×F×L×Exp	8	8.113	1.014	0.518	0.8408	
I×F×L×S/Exp	112	219.308	1.958			
<u>Total</u>	479	1482.481				

*Indicates effects determined to be statistically significant after applying the Greenhouse-Geisser correction ($p=0.10$).

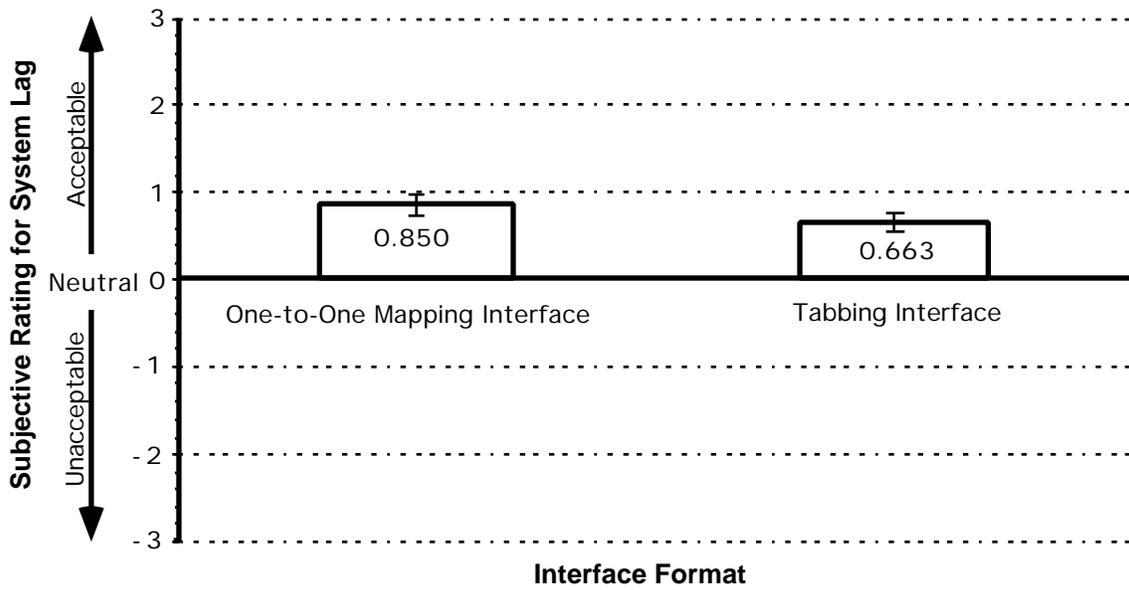


Figure 4.5 Effect of Interface Format on subjective ratings for System Lag. Error bars indicate ± 1 standard error of the mean (N=240).

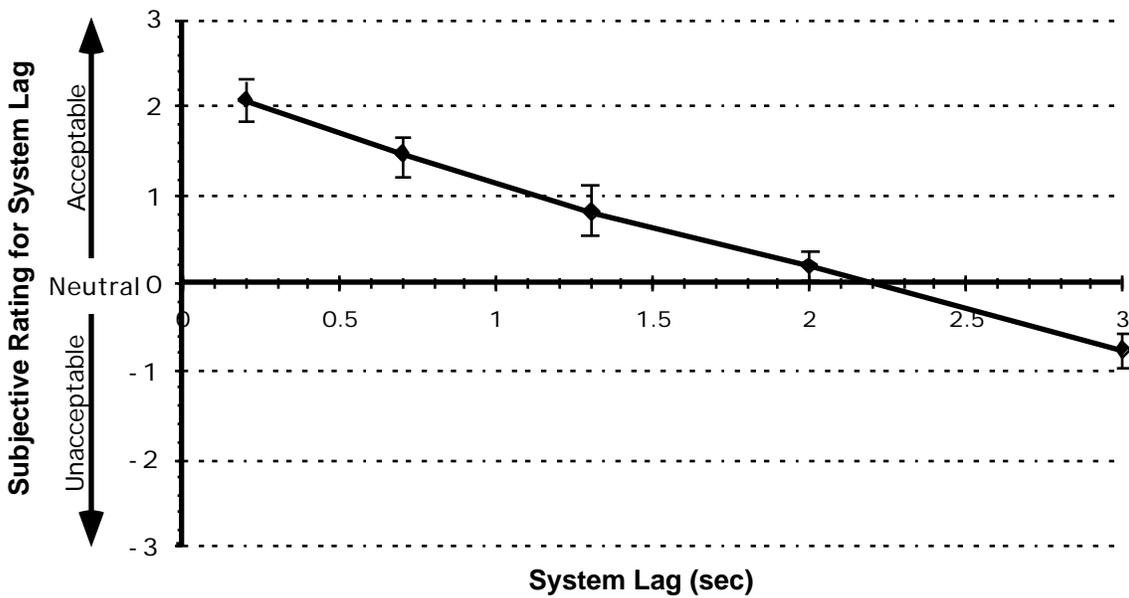


Figure 4.6 Effect of System Lag on subjective ratings for System Lag. Error bars indicate ± 1 standard error of the mean (N=96).

Table 4.5 Newman-Keuls Comparison of System Lag Subjective Ratings among System Lag Times.

System Lag (sec)	Mean	Group
0.2	2.0833	A
0.7	1.4583	B
1.3	0.8229	C
2.0	0.1979	D
3.0	-0.7813	E

Groups with the same letter are not significantly different at $p=0.10$

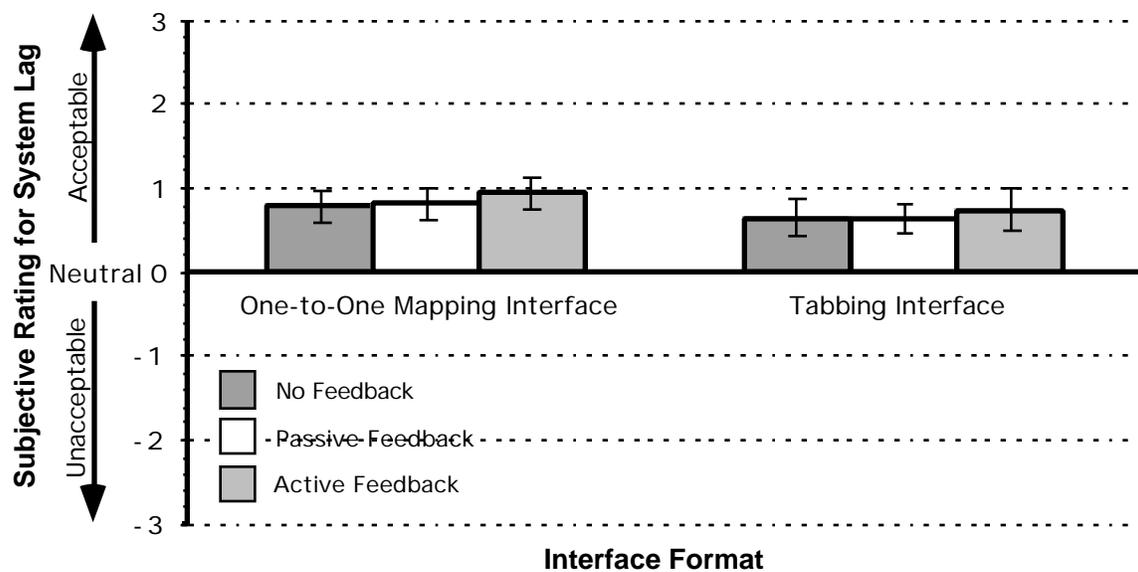


Figure 4.7 Effect of Interface Format \times Feedback Style interaction on subjective ratings for System Lag. Error bars indicate ± 1 standard error of the mean (N=80).

Table 4.6 Newman-Keuls Comparison of System Lag Subjective Ratings for Interface Format \times Feedback Style Interaction.

Interface Format	Feedback Style	Mean	Group
One-to-One Mapping	Active	0.9375	A
One-to-One Mapping	Passive	0.8250	A
One-to-One Mapping	No	0.7875	A
Tabbing	Active	0.7125	A
Tabbing	Passive	0.6500	A
Tabbing	No	0.6250	A

Groups with the same letter are not significantly different at $p=0.10$

Task Difficulty

Table 4.7 shows the ANOVA summary table for the subjective ratings on task difficulty as perceived by the participants. The main effects of Experience Level and Interface Format were significant, as were the Interface Format \times Feedback Style, Interface Format \times System Lag, and Interface Format \times Feedback Style \times System Lag interactions.

Figure 4.8 reveals that the inexperienced users rated the tasks to be 46% easier than did the experienced users. As seen in Figure 4.9, performing the tasks with the tabbing interface was rated 32% easier than performing them with the one-to-one mapping interface.

Figure 4.10 illustrates the effect of Interface Format \times Feedback Style interaction on task difficulty rating. On average, tasks with the tabbing interface and the passive feedback received the highest rank and tasks with the one-to-one mapping interface and the passive feedback received the lowest rank. Newman-Keuls post-hoc analysis (Table 4.8) reveals that tasks with the tabbing interface and any of the feedback styles were perceived to be significantly easier than the tasks performed with the one-to-one mapping interface and any of the feedback styles.

As demonstrated by Figure 4.11, tasks were perceived to be easier with the tabbing interface than with the one-to-one mapping interface at any level of system lag. Table 4.9 shows the Newman-Keuls post-hoc analysis that indicates the tasks performed with the tabbing interface and either the 0.2 or the 0.7 second system lag were rated significantly easier compared to the tasks performed with the one-to-one mapping interface and either the 1.3 or the 3.0 second system lag.

The effect of Interface Format \times Feedback Style \times System Lag interaction on task difficulty is presented in Figure 4.12. The results of the Newman-Keuls post-hoc analysis are shown in Table 4.10. When participants were presented with the one-to-one mapping interface; without feedback they preferred the 1.3 second system lag, with passive feedback they preferred the 0.2 second system lag, and with active feedback they preferred the 0.7 second system lag. The ratings for task difficulty were not determined to be significantly different when tabbing interface was used with different feedback styles and system lags.

Table 4.7 ANOVA Summary Table for Subjective Ratings on Task Difficulty.

Source	df	SS	MS	F	p	p_{G-G}
<u>Between-Subjects</u>						
Experience (Exp)	1	46.875	46.875	5.590	0.0330*	<0.10
Subjects/Experience (S/Exp)	14	117.392	8.385			
<u>Within-Subjects</u>						
Interface Format (I)	1	8.008	8.008	4.752	0.0468*	<0.10
I×Exp	1	0.000	0.000	0.000	1.0000	
I×S/Exp	14	23.592	1.685			
Feedback Style (F)	2	0.129	0.065	0.027	0.9734	
F×Exp	2	0.913	0.456	0.192	0.8264	
F×S/Exp	28	66.558	2.377			
System Lag (L)	4	14.550	3.638	1.903	0.1227	
L×Exp	4	10.250	2.562	1.340	0.2655	
L×S/Exp	56	107.067	1.912			
I×F	2	32.879	16.440	11.505	0.0002*	<0.10
I×F×Exp	2	3.912	1.956	1.369	0.2709	
I×F×S/Exp	28	40.008	1.429			
I×L	4	57.783	14.446	5.539	0.0008*	<0.10
I×L×Exp	4	1.250	0.312	0.120	0.9748	
I×L×S/Exp	56	146.033	2.608			
F×L	8	29.850	3.731	2.837	0.0066	>0.10
F×L×Exp	8	8.650	1.081	0.822	0.5848	
F×L×S/Exp	112	147.233	1.315			
I×F×L	8	66.392	8.299	3.835	0.0005*	<0.10
I×F×L×Exp	8	11.775	1.472	0.680	0.7083	
I×F×L×S/Exp	112	242.367	2.164			
<u>Total</u>	479	1183.467				

*Indicates effects determined to be statistically significant after applying the Greenhouse-Geisser correction (p=0.10).

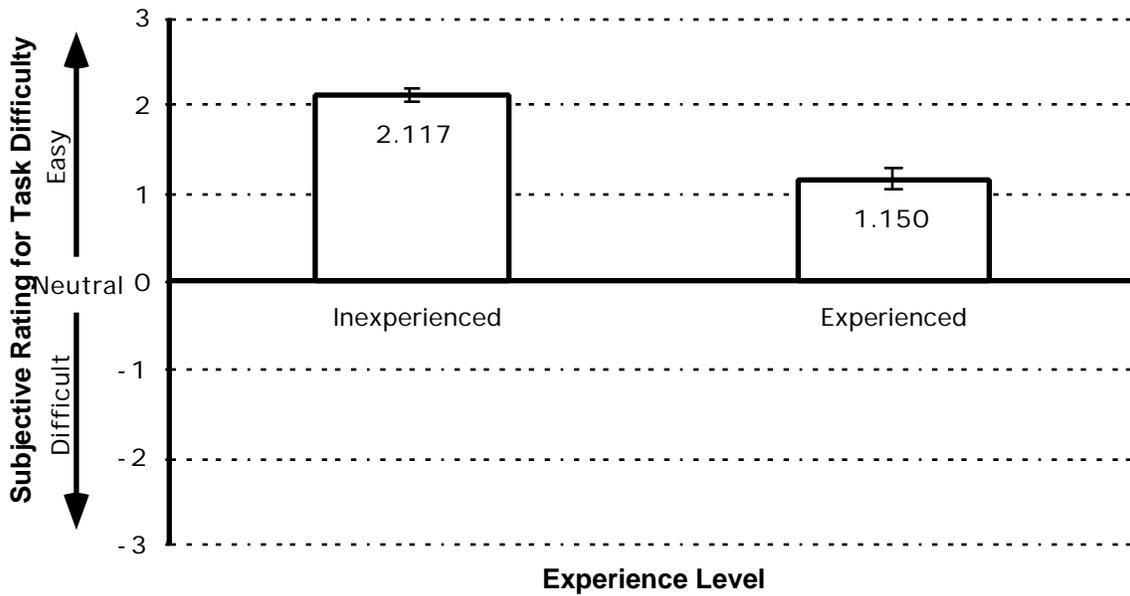


Figure 4.8 Effect of Experience Level on subjective ratings for task difficulty. Error bars indicate ± 1 standard error of the mean (N=240).

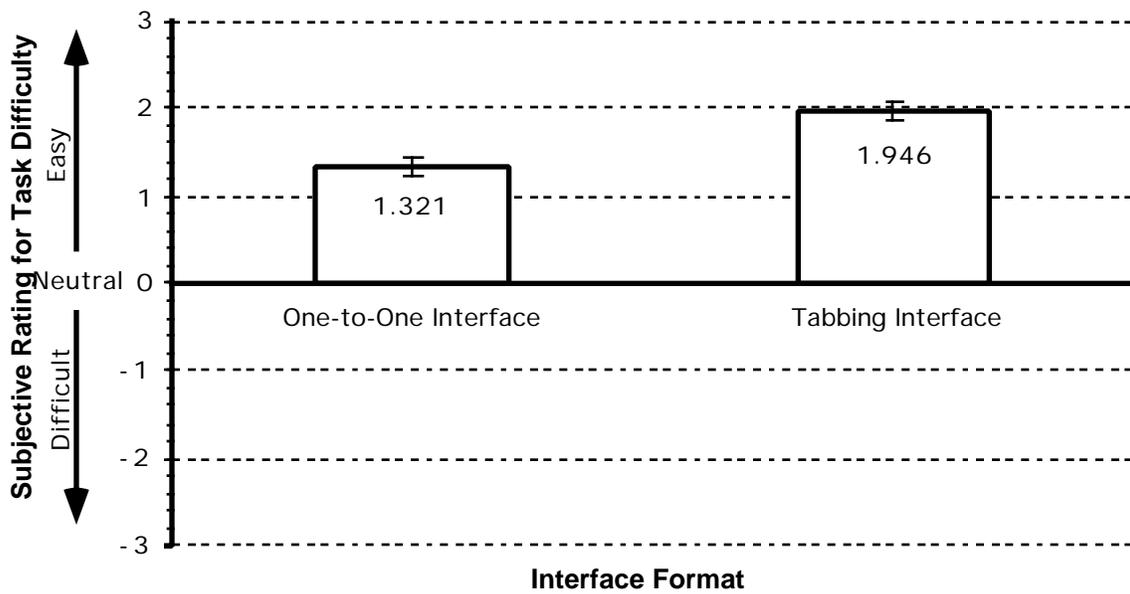


Figure 4.9 Effect of Interface Format on subjective ratings for task difficulty. Error bars indicate ± 1 standard error of the mean (N=240).

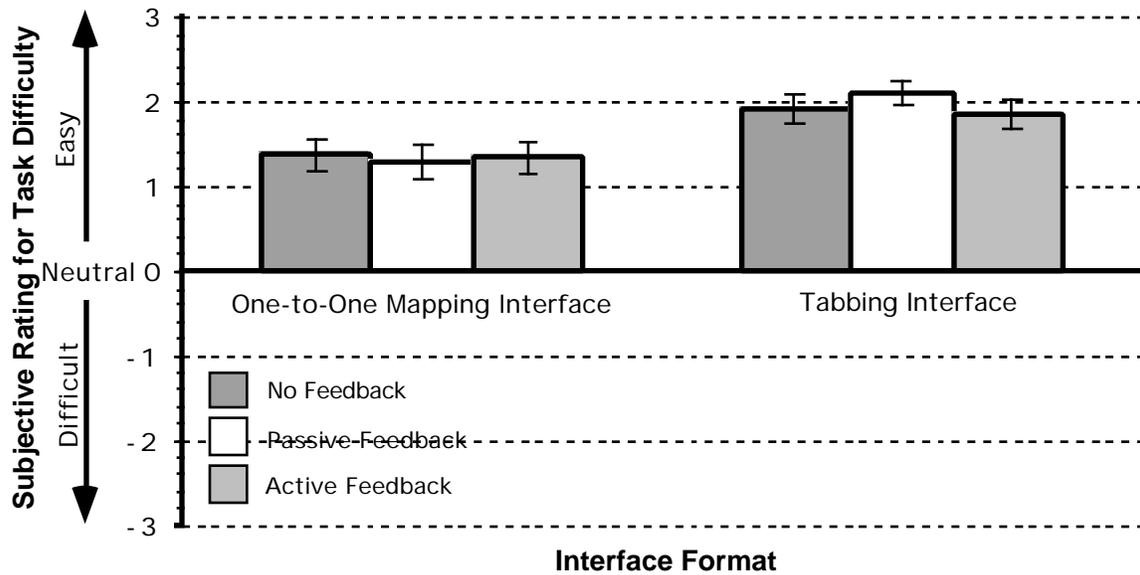


Figure 4.10 Effect of Interface Format \times Feedback Style interaction on subjective ratings for task difficulty. Error bars indicate ± 1 standard error of the mean (N=80).

Table 4.8 Newman-Keuls Comparison of Task Difficulty Subjective Ratings for Interface Format \times Feedback Style Interaction.

Interface Format	Feedback Style	Mean	Group
Tabbing	Passive	2.0875	A
Tabbing	No	1.9125	A
Tabbing	Active	1.8375	A
One-to-One Mapping	No	1.3625	B
One-to-One Mapping	Active	1.3250	B
One-to-One Mapping	Passive	1.2750	B

Groups with the same letter are not significantly different at $p=0.10$

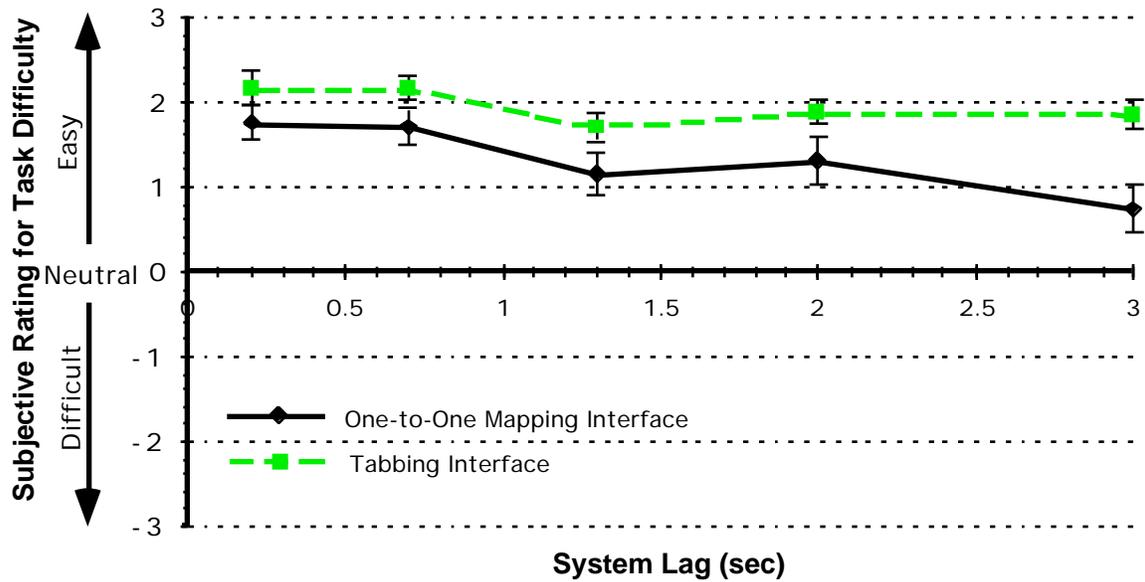


Figure 4.11 Effect of Interface Format \times System Lag interaction on subjective ratings for task difficulty. Error bars indicate ± 1 standard error of the mean (N=48).

Table 4.9 Newman-Keuls Comparison of Task Difficulty Subjective Ratings for Interface Format \times System Lag Interaction.

Interface Format	System Lag (sec)	Mean	Group
Tabbing	0.2	2.1458	A
Tabbing	0.7	2.1458	A
Tabbing	2.0	1.8750	A B
Tabbing	3.0	1.8542	A B
One-to-One Mapping	0.2	1.7500	A B
Tabbing	1.3	1.7083	A B
One-to-One Mapping	0.7	1.6875	A B
One-to-One Mapping	2.0	1.2917	A B C
One-to-One Mapping	1.3	1.1458	B C
One-to-One Mapping	3.0	0.7292	C

Groups with the same letter are not significantly different at $p=0.10$

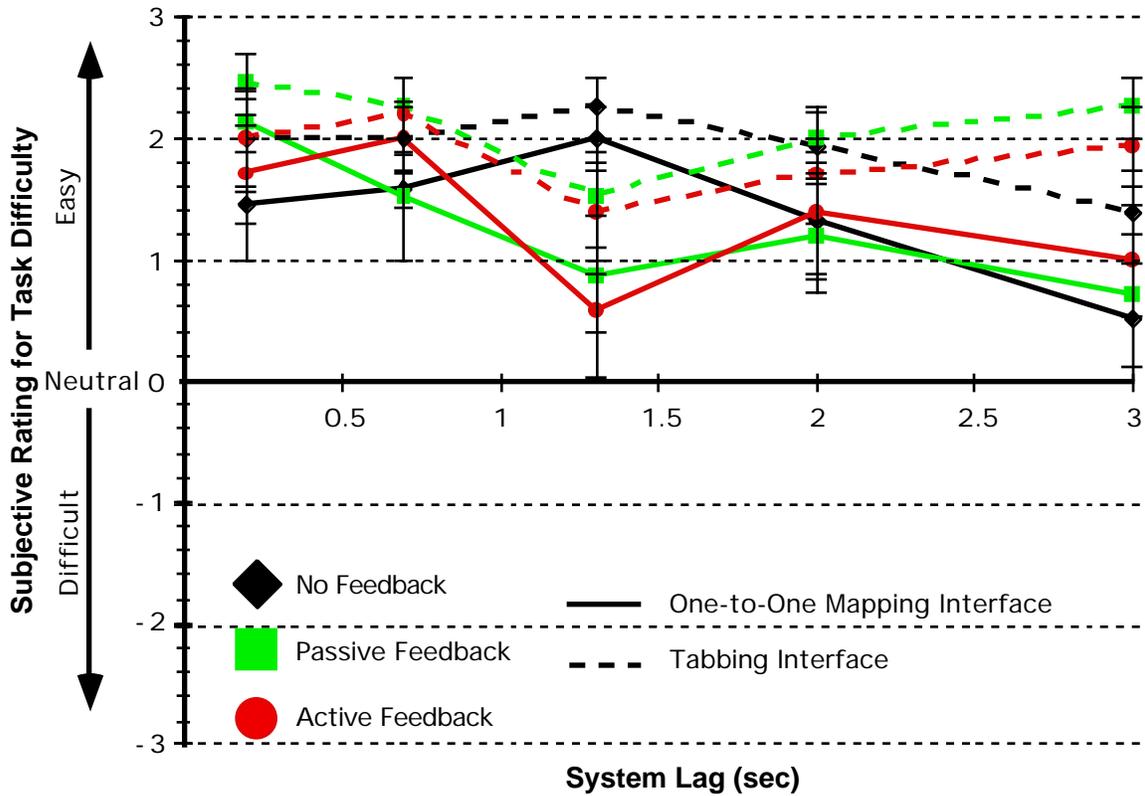


Figure 4.12 Effect of Interface Format × Feedback Style × System Lag interaction on subjective ratings for task difficulty. Error bars indicate ± 1 standard error of the mean (N=16).

Table 4.10 Newman-Keuls Comparison of Task Difficulty Subjective Ratings for Interface Format × Feedback Style × System Lag Interaction.

One-to-One Mapping Interface with No Feedback			Tabbing Interface with No Feedback		
System Lag (sec)	Mean	Group	System Lag (sec)	Mean	Group
1.3	2.0000	A	1.3	2.2500	A
0.7	1.5625	A B	0.2	2.0000	A
0.2	1.4375	A B	0.7	2.0000	A
2.0	1.3125	A B	2.0	1.9375	A
3.0	0.5000	B	3.0	1.3750	A

One-to-One Mapping Interface with Passive Feedback			Tabbing Interface with Passive Feedback		
System Lag (sec)	Mean	Group	System Lag (sec)	Mean	Group
0.2	2.1250	A	0.2	2.4375	A
0.7	1.5000	A B	0.7	2.2500	A
2.0	1.1875	A B	3.0	2.2500	A
1.3	0.8750	B	2.0	2.0000	A
3.0	0.6875	B	1.3	1.5000	A

One-to-One Mapping Interface with Active Feedback			Tabbing Interface with Active Feedback		
System Lag (sec)	Mean	Group	System Lag (sec)	Mean	Group
0.7	2.0000	A	0.7	2.1875	A
0.2	1.6875	A B	0.2	2.0000	A
2.0	1.3750	A B	3.0	1.9375	A
3.0	1.0000	A B	2.0	1.6875	A
1.3	0.5625	B	1.3	1.3750	A

Groups with the same letter are not significantly different at p=0.10

Feedback Style

The ANOVA summary table for the subjective ratings on Feedback Style is presented in Table 4.11. The main effects of Interface Format and Feedback Style were significant, as well as the Interface Format × Feedback Style and the Interface Format × System Lag interactions.

Figure 4.13 illustrates that on average, participants rated the acceptability of feedback styles 9% higher when they used the tabbing interface than when they used the one-to-one mapping interface. Passive feedback was rated 85% better than no feedback, while active feedback was considered 38% better than passive feedback (Figure 4.14). The Newman-Keuls

post-hoc comparison (Table 4.12) reveals that feedback styles were rated significantly different, with active feedback being the highest and no feedback the lowest.

Analysis of the interaction between interface format and feedback style reveals similar findings to the previous results on feedback style ratings (Figure 4.15). According to the Newman-Keuls post-hoc analysis, feedback styles were rated significantly different within each interface style; active feedback received the highest rating and no feedback the lowest (Table 4.13). The interaction between interface format and system lag is presented in Figure 4.16. As seen in Table 4.14, the Newman-Keuls test revealed no significant differences in feedback style ratings for this interaction.

Table 4.11 ANOVA Summary Table for Subjective Ratings on Feedback Style.

Source	df	SS	MS	F	p	p _{G-G}
<u>Between-Subjects</u>						
Experience (Exp)	1	1.408	1.408	0.134	0.7198	
Subjects/Experience (S/Exp)	14	146.917	10.494			
<u>Within-Subjects</u>						
Interface Format (I)	1	163.333	163.333	30.236	0.0001*	<0.10
I×Exp	1	0.300	0.300	0.056	0.8164	
I×S/Exp	14	75.633	5.402			
Feedback Style (F)	2	68.517	34.258	30.506	0.0000*	<0.10
F×Exp	2	1.317	0.658	0.586	0.5632	
F×S/Exp	28	31.433	1.123			
System Lag (L)	4	4.429	1.107	1.159	0.3387	
L×Exp	4	3.071	0.768	0.804	0.5278	
L×S/Exp	56	53.500	0.955			
I×F	2	76.517	38.258	21.349	0.0000*	<0.10
I×F×Exp	2	2.450	1.225	0.684	0.5128	
I×F×S/Exp	28	50.167	1.792			
I×L	4	14.854	3.714	4.985	0.0017*	<0.10
I×L×Exp	4	1.513	0.378	0.507	0.7307	
I×L×S/Exp	56	41.700	0.745			
F×L	8	5.796	0.724	0.555	0.8125	
F×L×Exp	8	6.579	0.822	0.630	0.7511	
F×L×S/Exp	112	146.025	1.304			
I×F×L	8	48.171	6.021	2.325	0.0239	>0.10
I×F×L×Exp	8	4.238	0.530	0.205	0.9895	
I×F×L×S/Exp	112	290.125	2.590			
<u>Total</u>	479	1237.992				

*Indicates effects determined to be statistically significant after applying the Greenhouse-Geisser correction (p=0.10).

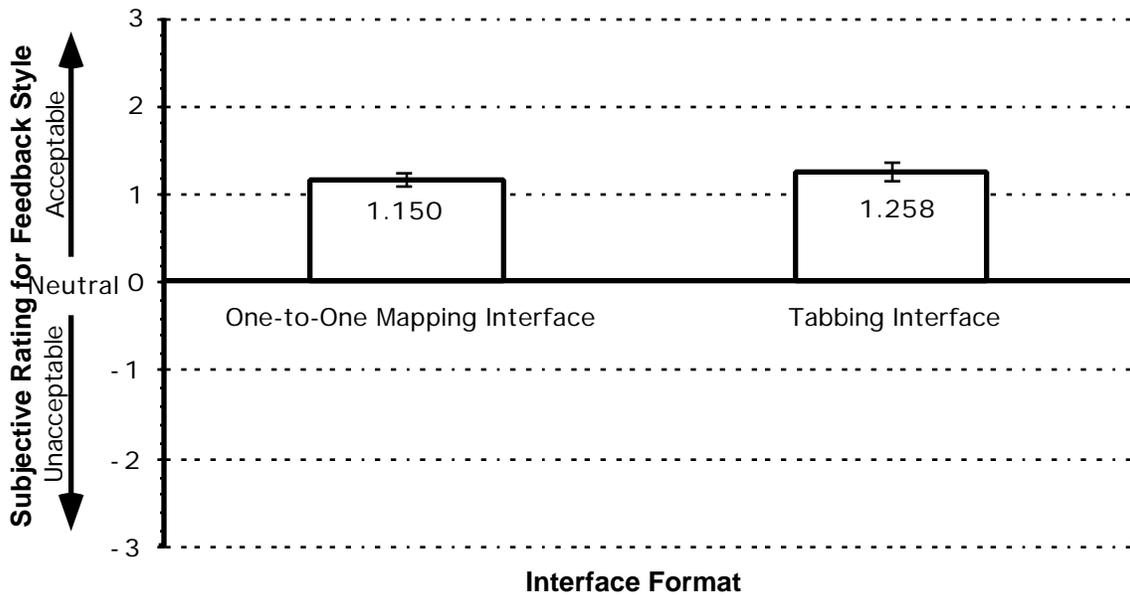


Figure 4.13 Effect of Interface Format on subjective ratings for Feedback Style. Error bars indicate ± 1 standard error of the mean (N=240).

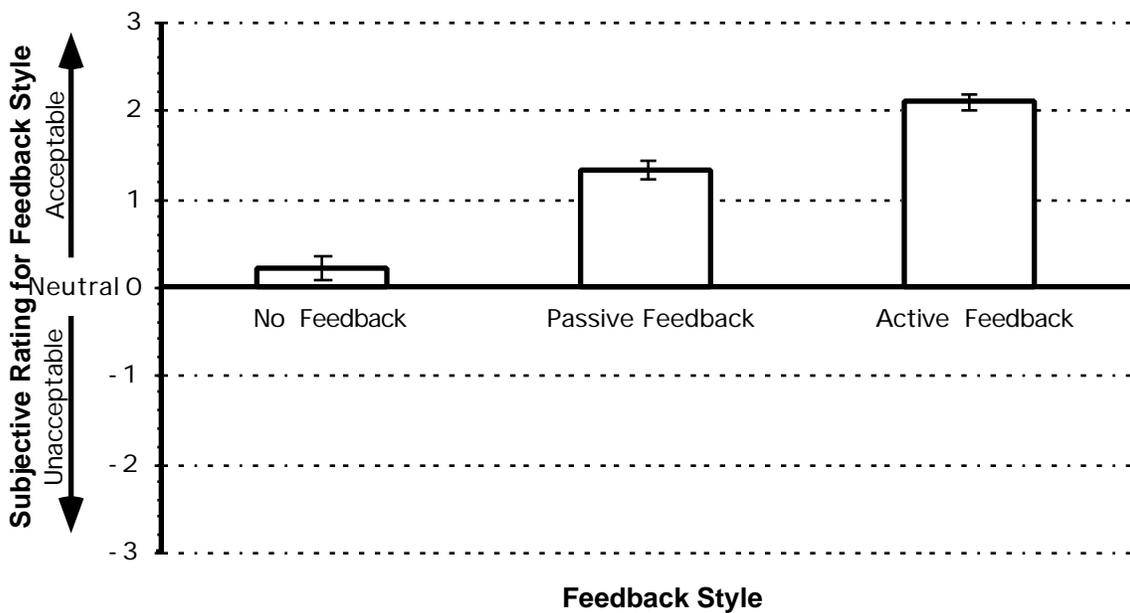


Figure 4.14 Effect of Feedback Style on subjective ratings for Feedback Style. Error bars indicate ± 1 standard error of the mean (N=160).

Table 4.12 Newman-Keuls Comparison of Feedback Style Subjective Ratings among Feedback Styles.

Feedback Styles	Mean	Group
Active	2.1063	A
Passive	1.3063	B
No	0.2000	C

Groups with the same letter are not significantly different at $p=0.10$

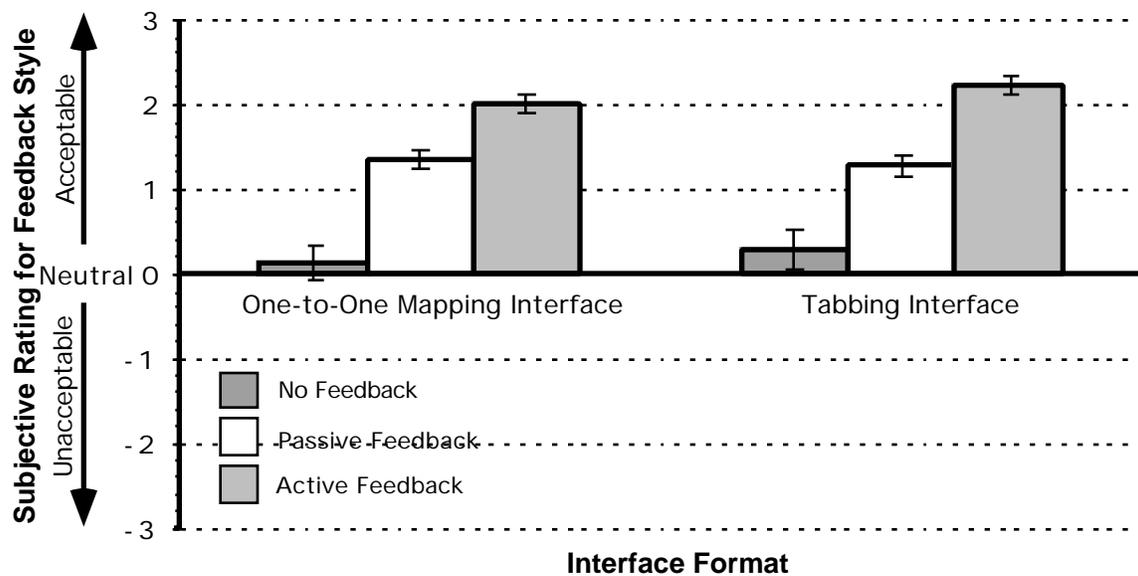


Figure 4.15 Effect of Interface Format \times Feedback Style interaction on subjective ratings for Feedback Style. Error bars indicate ± 1 standard error of the mean (N=80).

Table 4.13 Newman-Keuls Comparison of Feedback Style Subjective Ratings for Interface Format × Feedback Style Interaction.

Interface Format	Feedback Style	Mean	Group
Tabbing	Active	2.2125	A
One-to-One Mapping	Active	2.0000	A
One-to-One Mapping	Passive	1.3250	B
Tabbing	Passive	1.2875	B
Tabbing	No	0.2750	C
One-to-One Mapping	No	0.1250	C

Groups with the same letter are not significantly different at $p=0.10$

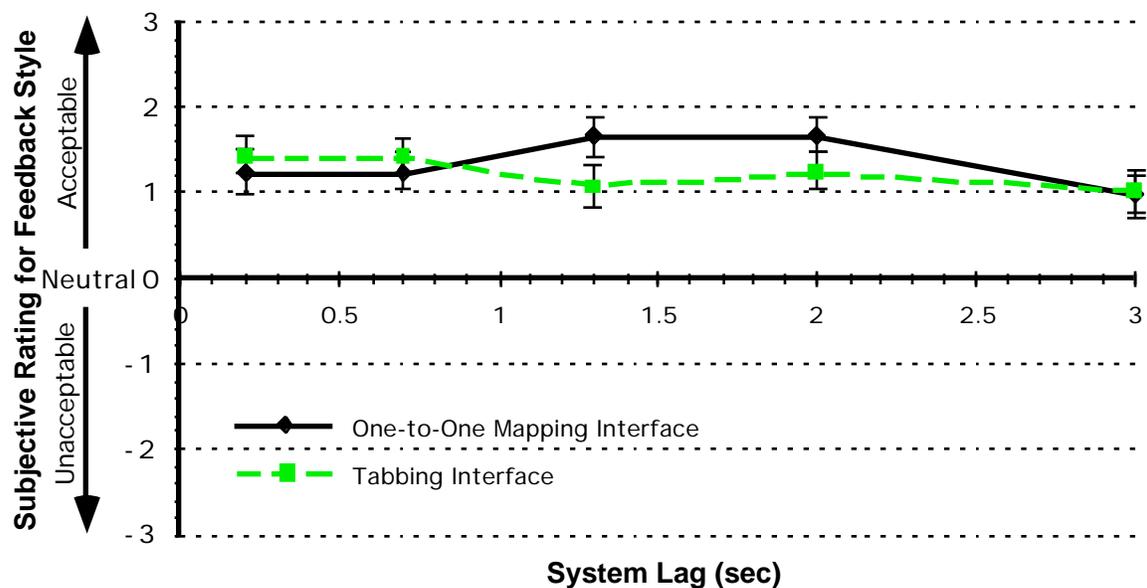


Figure 4.16 Effect of Interface × System Lag interaction on subjective ratings for Feedback Style. Error bars indicate ± 1 standard error of the mean ($N=48$).

Table 4.14 Newman-Keuls Comparison of Feedback Style Subjective Ratings for Interface Format x System Lag Interaction.

Interface Format	System Lag (sec)	Mean	Group
Tabbing	0.2	1.4375	A
Tabbing	0.7	1.4375	A
Tabbing	2.0	1.2500	A
One-to-One Mapping	0.2	1.2292	A
One-to-One Mapping	0.7	1.2292	A
One-to-One Mapping	1.3	1.1667	A
One-to-One Mapping	2.0	1.1667	A
Tabbing	3.0	1.1042	A
Tabbing	1.3	1.0625	A
One-to-One Mapping	3.0	0.9583	A

Groups with the same letter are not significantly different at $p=0.10$

Usability

The ANOVA summary table for the usability ratings is shown in Table 4.15. The main effects of Experience Level, Feedback Style, and System Lag were significant. The interactions that were determined to be significant were as follows: Interface Format x Feedback Style, Interface Format x System Lag, and Interface Format x Feedback Style x System Lag.

As demonstrated in Figure 4.17, inexperienced users' usability ratings were 47% higher than the experienced users' ratings. The effect of feedback style on usability is illustrated in Figure 4.18. Newman-Keuls post-hoc analysis (Table 4.16) reveals that the three feedback styles resulted in significantly different usability ratings, with active feedback being the highest and no feedback the lowest. The effect of system lag on usability is presented in Figure 4.19. The Newman-Keuls post-hoc analysis (Table 4.17) indicated that system lags of 0.2 and 0.7 seconds resulted in significantly better usability ratings than did the 1.3 and 3.0 second system lags, which in turn were rated significantly better than the 3.0 second system lag.

Figure 4.20 shows the interaction between Interface Format and Feedback Style. On average, participants rated the usability of the tabbing interface with the active feedback the highest and the usability of the one-to-one mapping interface with no feedback the lowest. The results of the Newman-Keuls post-hoc analysis presented in Table 4.18 reveals that the tabbing interface with all the feedback styles and the one-to-one mapping interface with active feedback were rated significantly higher in usability ratings than the one-to-one mapping interface with no feedback.

Figure 4.21 shows the interaction between interface format and system lag. 0.2 second lag with the tabbing interface was rated the highest and the 3.0 second lag with the one-to-one mapping interface was rated the lowest. According to the Newman-Keuls post-hoc analysis (Table 4.19), the tabbing interface and 0.2 second lag received significantly better usability ratings than did the other combinations of interface format with 1.3, 2.0, and 3.0 second system lags.

Figure 4.22 illustrates the interaction among interface format, feedback style, and system lag. The Newman-Keuls post-hoc analysis is shown in Table 4.20. When participants were presented with the one-to-one mapping interface and no feedback they rated the 3.0 second system lag significantly lower than the other ones. However, when active feedback was used, 0.2 second system lag was rated significantly higher than the 3.0 second system lag. No significant differences were found in the usability ratings for the rest of the interface format, feedback style, and system lag combinations.

Subjective ratings for the two Interface Formats

After performing fifteen tasks with each interface format, participants were asked to evaluate the interface formats by completing Likert-type scales (Appendix C). The results of these evaluations are presented in Figure 4.23. On average, experienced users' ratings were lower than the inexperienced users' ratings ($p=0.000$) and the tabbing interface received higher ratings than did the one-to-one mapping interface ($p=0.001$). Participants rated the number of topic choices on the screen to be more acceptable for the tabbing interface than for the one-to-one mapping interface ($p=0.012$); inexperienced users felt the same way ($p=0.036$). For the experienced users the topics in the tree structure were easier to follow with the tabbing interface than they were with the one-to-one mapping interface ($p=0.019$). Inexperienced users gave significantly higher ratings than did the experienced users for several issues concerning the one-to-one mapping interface: the ease of following topics in the tree structure ($p=0.041$), the ease of navigating through the tree structure by using the nine buttons ($p=0.033$), and the ease of adapting to the change in button functions ($p=0.095$).

Participants' preferences

Each participant was asked to indicate their preference of feedback style and interface format at the end of the experimental session (Appendix C). 37.5% of the inexperienced users and 25% of the experienced users preferred the one-to-one mapping interface. All the participants preferred the active feedback.

Table 4.15 ANOVA Summary Table for Subjective Ratings on the Usability of the System.

Source	df	SS	MS	F	p	p_{G-G}
<u>Between-Subjects</u>						
Experience (Exp)	1	10.800	10.800	12.424	0.0034*	<0.10
Subjects/Experience (S/Exp)	14	121.700	8.693			
<u>Within-Subjects</u>						
Interface Format (I)	1	0.075	0.075	0.125	0.7289	
I×Exp	1	1.008	1.008	1.683	0.2155	
I×S/Exp	14	8.383	0.599			
Feedback Style (F)	2	5.254	2.627	4.856	0.0155*	<0.10
F×Exp	2	0.262	0.131	0.242	0.7867	
F×S/Exp	28	15.150	0.541			
System Lag (L)	4	19.133	4.783	5.735	0.0006*	<0.10
L×Exp	4	3.617	0.904	1.084	0.3733	
L×S/Exp	56	46.717	0.834			
I×F	2	74.512	37.256	25.588	0.0000*	<0.10
I×F×Exp	2	1.254	0.627	0.431	0.6541	
I×F×S/Exp	28	40.767	1.456			
I×L	4	22.383	5.596	5.454	0.0009*	<0.10
I×L×Exp	4	1.367	0.342	0.333	0.8546	
I×L×S/Exp	56	57.450	1.026			
F×L	8	12.892	1.611	2.867	0.0061	>0.10
F×L×Exp	8	3.508	0.439	0.781	0.6203	
F×L×S/Exp	112	62.933	0.562			
I×F×L	8	100.092	12.511	11.859	0.0000*	<0.10
I×F×L×Exp	8	8.558	1.070	1.014	0.4297	
I×F×L×S/Exp	112	118.150	1.055			
<u>Total</u>	479	735.967				

*Indicates effects determined to be statistically significant after applying the Greenhouse-Geisser correction (p=0.10).

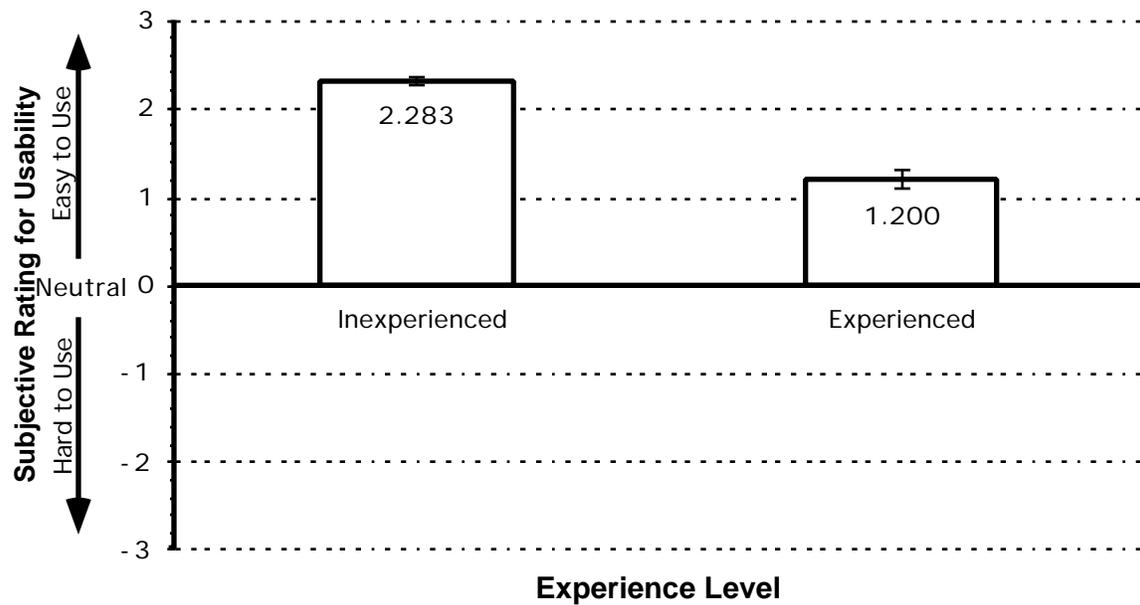


Figure 4.17 Effect of Experience Level on subjective ratings for usability of the system. Error bars indicate ± 1 standard error of the mean (N=240).

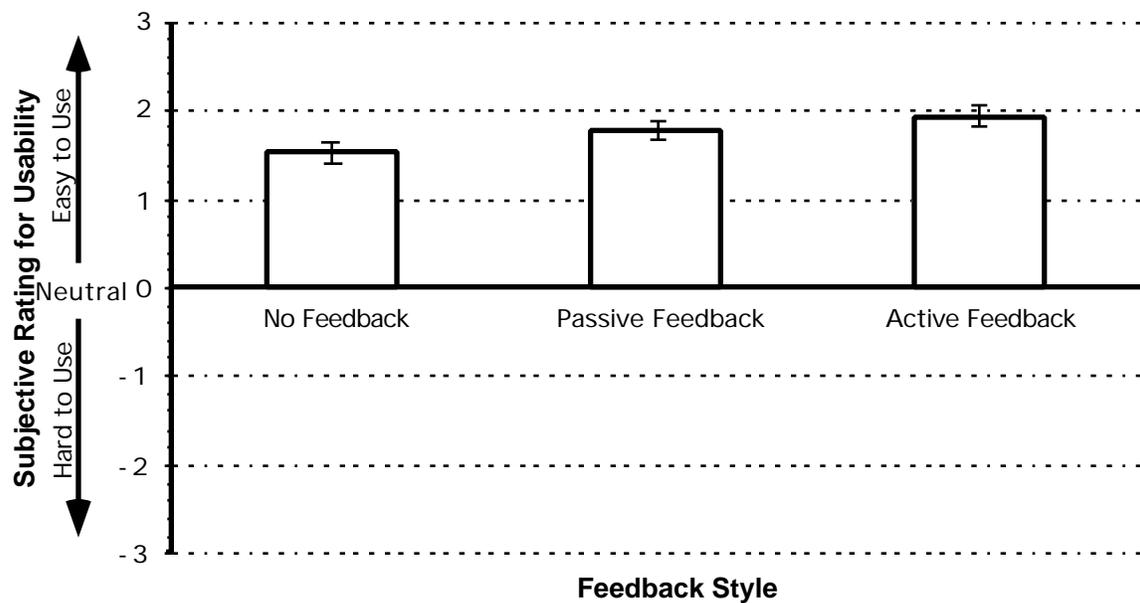


Figure 4.18 Effect of Feedback Style on subjective ratings for usability of the system. Error bars indicate ± 1 standard error of the mean (N=160).

Table 4.16 Newman-Keuls Comparison of Usability Subjective Ratings among Feedback Styles.

Feedback Styles	Mean	Group
Active	1.9120	A
Passive	1.7610	B
No	1.5283	C

Groups with the same letter are not significantly different at $p=0.10$

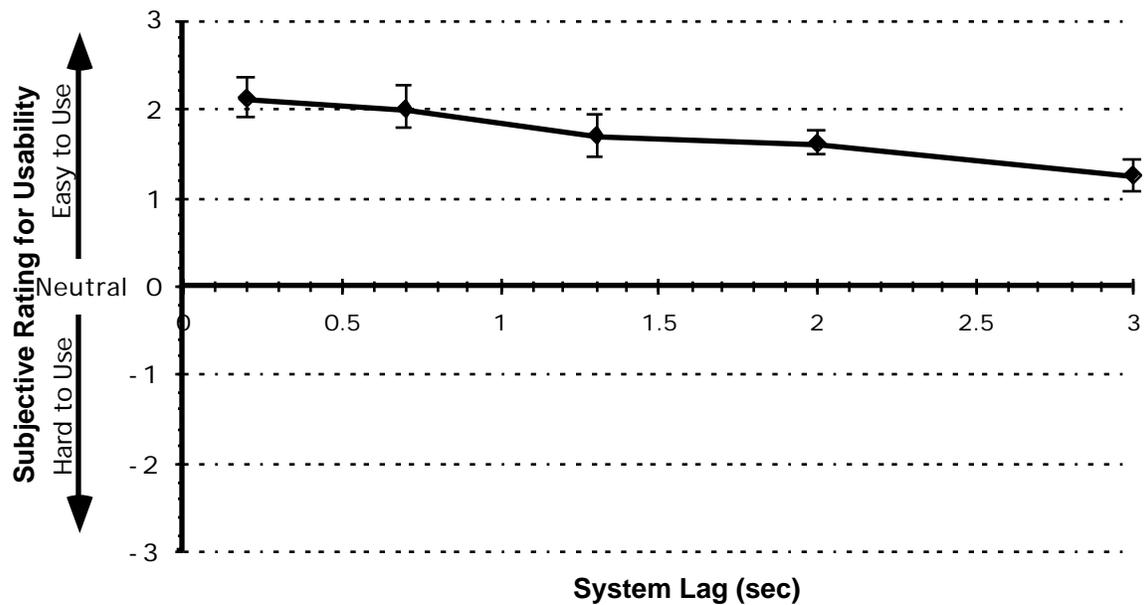


Figure 4.19 Effect of System Lag on subjective ratings for usability of the system. Error bars indicate ± 1 standard error of the mean (N=96).

Table 4.17 Newman-Keuls Comparison of Usability Subjective Ratings among System Lag Times.

System Lag (sec)	Mean	Group
0.2	2.1354	A
0.7	2.0000	A
1.3	1.6979	B
2.0	1.6250	B
3.0	1.2500	C

Groups with the same letter are not significantly different at $p=0.10$

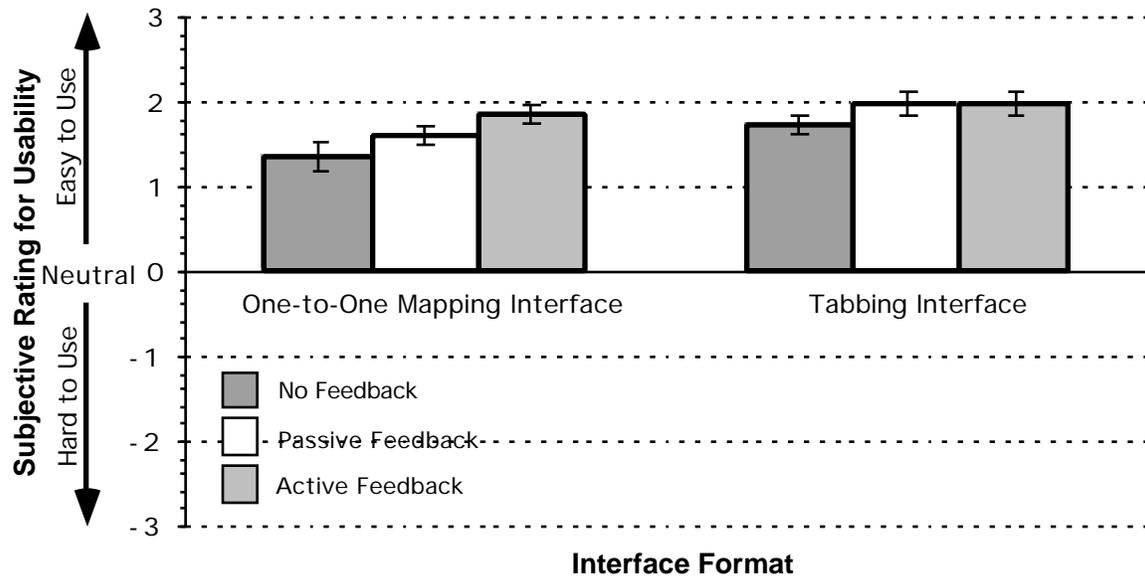


Figure 4.20 Effect of Interface Format × Feedback Style interaction on subjective ratings for usability of the system. Error bars indicate ± 1 standard error of the mean (N=80).

Table 4.18 Newman-Keuls Comparison of Usability Subjective Ratings for Interface Format × Feedback Style Interaction.

Interface Format	Feedback Style	Mean	Group
Tabbing	Active	1.9750	A
Tabbing	Passive	1.9625	A
One-to-One Mapping	Active	1.8625	A
Tabbing	No	1.7375	A
One-to-One Mapping	Passive	1.5750	A B
One-to-One Mapping	No	1.3375	B

Groups with the same letter are not significantly different at $p=0.10$

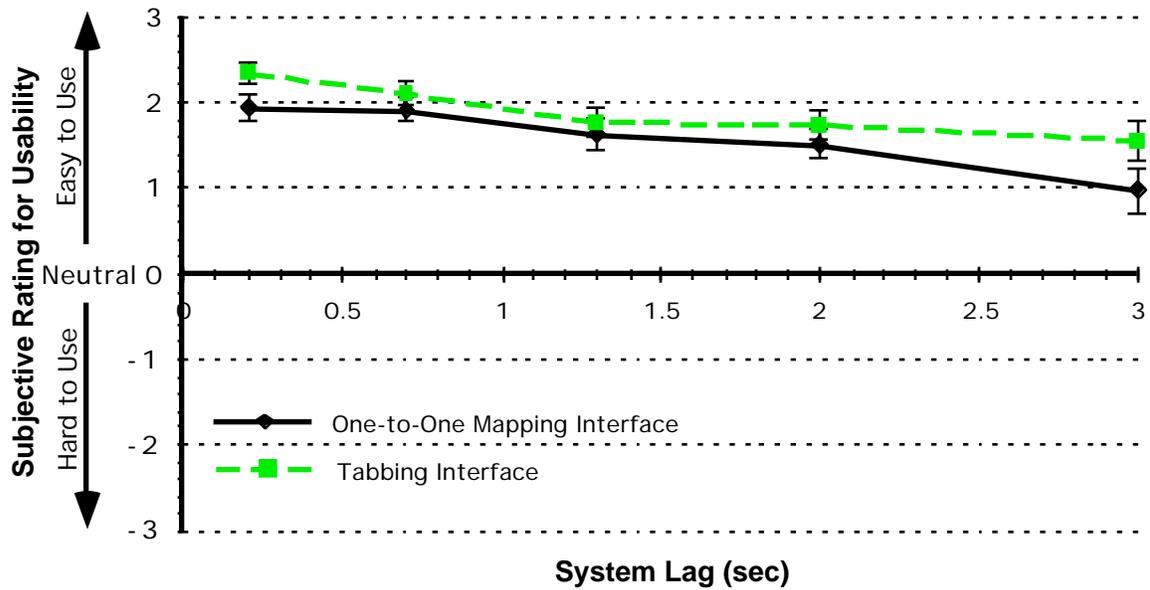


Figure 4.21 Effect of Interface Format \times System Lag interaction on subjective ratings for usability of the system. Error bars indicate ± 1 standard error of the mean (N=48).

Table 4.19 Newman-Keuls Comparison of Usability Subjective Ratings for Interface Format \times System Lag Interaction.

Interface Format	System Lag (sec)	Mean	Group
Tabbing	0.2	2.3333	A
Tabbing	0.7	2.1042	A B
One-to-One Mapping	0.2	1.9375	A B
One-to-One Mapping	0.7	1.8958	A B
Tabbing	1.3	1.7500	B
Tabbing	2.0	1.7292	B
One-to-One Mapping	1.3	1.6458	B
Tabbing	3.0	1.5417	B
One-to-One Mapping	2.0	1.5208	B
One-to-One Mapping	3.0	0.9583	C

Groups with the same letter are not significantly different at $p=0.10$

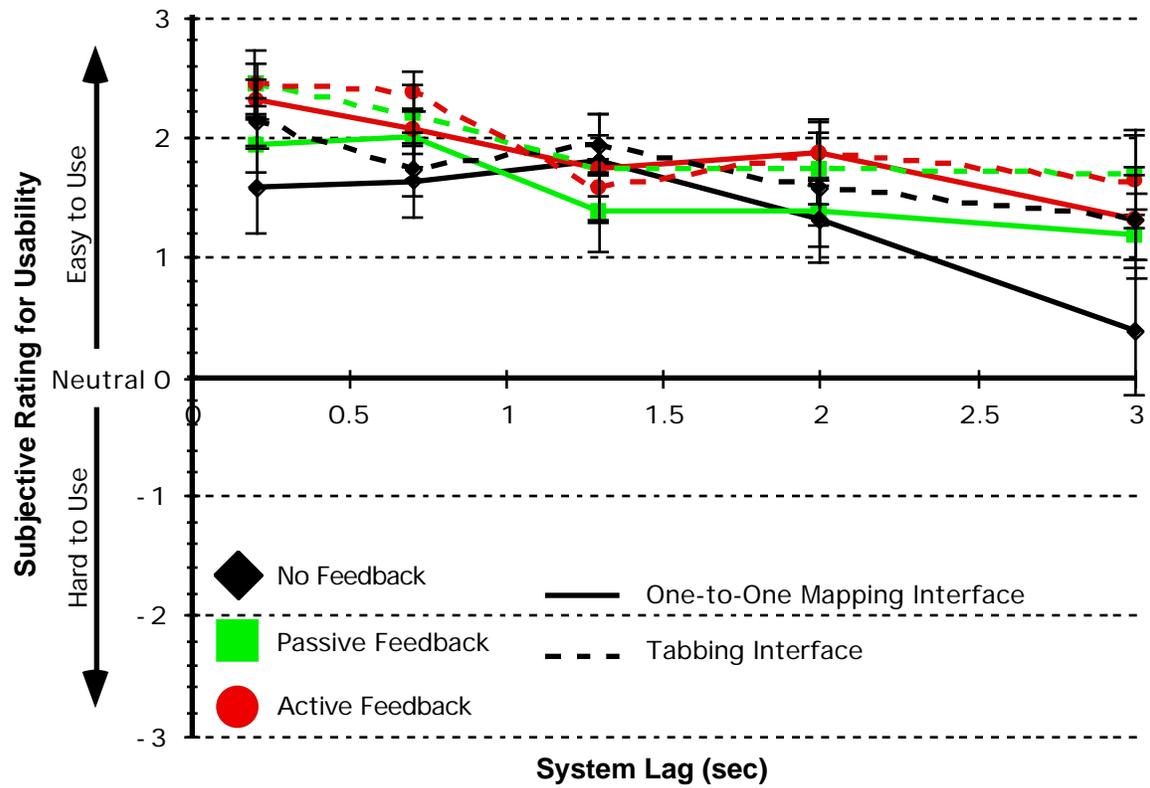


Figure 4.22 Effect of Interface Format \times Feedback Style \times System Lag interaction on subjective ratings for usability of the system. Error bars indicate ± 1 standard error of the mean (N=16).

Table 4.20 Newman-Keuls Comparison of Usability Subjective Ratings for Interface Format × Feedback Style × System Lag Interaction.

One-to-One Mapping Interface
with No Feedback

System Lag (sec)	Mean	Group
1.3	1.8125	A
0.7	1.6250	A
0.2	1.5625	A
2.0	1.3125	A
3.0	0.3750	B

Tabbing Interface with No Feedback

System Lag (sec)	Mean	Group
0.2	2.1250	A
1.3	1.9375	A
0.7	1.7500	A
2.0	1.5625	A
3.0	1.3125	A

One-to-One Mapping Interface
with Passive Feedback

System Lag (sec)	Mean	Group
0.7	2.0000	A
0.2	1.9375	A
1.3	1.3750	A
2.0	1.3750	A
3.0	1.1875	A

Tabbing Interface with Passive Feedback

System Lag (sec)	Mean	Group
0.2	2.4375	A
0.7	2.1875	A
2.0	1.7500	A
1.3	1.7500	A
3.0	1.6875	A

One-to-One Mapping Interface
with Active Feedback

System Lag (sec)	Mean	Group
0.2	2.3125	A
0.7	2.0625	A B
2.0	1.8750	A B
1.3	1.7500	A B
3.0	1.3125	B

Tabbing Interface with Active Feedback

System Lag (sec)	Mean	Group
0.2	2.4375	A
0.7	2.3750	A
2.0	1.8750	A
3.0	1.6250	A
1.3	1.5625	A

Groups with the same letter are not significantly different at p=0.10

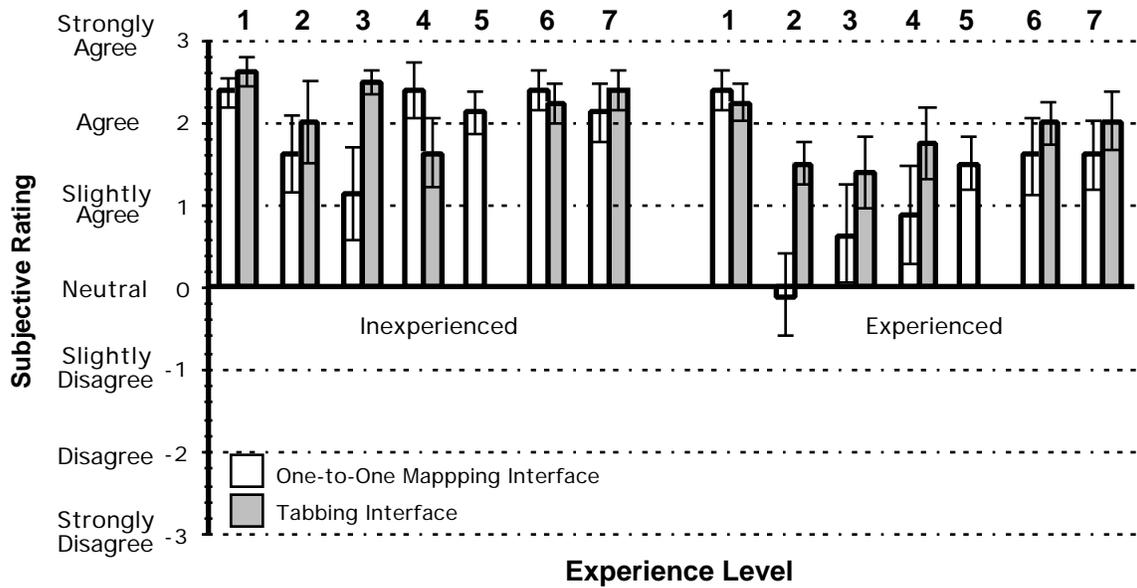


Figure 4.23 Results of the subjective ratings for each Interface Format. Error bars indicate ± 1 standard error of the mean (N=8) (Each column represents a question - refer to Appendix C for the questions).

CHAPTER 5: DISCUSSION

System Lag

Although system lag ratings showed 0.2 seconds to be significantly better than the other system lag times, there was no significant difference on usability ratings for the 0.2 and 0.7 second system lags. Therefore, 0.2 second system lag would be preferable. However, if necessary, a lag of 0.7 seconds would be acceptable.

Feedback Style

Both feedback ratings and usability ratings revealed active feedback to be better than passive feedback, which in turn was better than lack of feedback. Previous studies have shown that variable response times increase the need for active feedback (Planas and Treurniet, 1988; Brown, 1988). Therefore, it is advisable for systems accessing the Internet to have active feedback, by which a dynamic indicator is used to inform users about how much of their request has been processed.

Interface Format

The tree structure in the one-to-one mapping interface had up to eight topics per level and was up to six levels deep. The tabbing interface, on the other hand, had up to eighteen topics per level and was up to four levels deep. As a result, the one-to-one mapping interface was narrower, but deeper than the tabbing interface.

Participants' Likert-type scale ratings were higher for the tabbing interface than for the one-to-one mapping interface. Task completion time was 14% shorter when participants used the tabbing interface since there were more topic choices per level. For the same reason, participants found it easier to complete tasks and therefore, committed less errors. There was a 53% reduction in errors when the tabbing interface was used. These results suggest that narrow and deep tree structures may hinder accuracy and speed. Some other researchers such as Snowberry, et. al. (1983), Kiger (1984), Landauer and Nachbar (1985), and Paap and Roske-Hofstrand (1986) have reported similar effects of breadth and depth issues in tree structures. As observed by some participants, one drawback of broad tree structures may be the high number of categories per level, creating the need for extensive visual search for finding the desired topic. Similar findings have been reported by Paap and Roske-Hofstrand (1986) where they named the

phenomenon “crowding”. They determined that as users gain experience, the breadth and depth trade-off on processing time becomes insignificant. This finding may suggest that if the indexing structure does not change frequently, some of the disadvantages of one-to-one mapping interface may diminish. However, due to the nature of the Internet, indexing structures based on the WWW are likely to be dynamic.

Participants criticized the *cancel* button in the one-to-one mapping interface indicating that it was time consuming and frustrating to go back to a previous screen or go to the beginning of the index. Long system lags were more noticeable in the tabbing interface than in the one-to-one mapping interface since by design more button presses were expected in the tabbing interface. Therefore, system lag should be a key factor in choosing the appropriate interface format for the browsing system.

62.5% of inexperienced users and 75% of experienced users preferred the tabbing interface over the one-to-one mapping interface, averaging to 69% of the participants. All the participants who preferred the one-to-one mapping interface were exposed to the tabbing interface first, which may support the previously mentioned theory that the disadvantages of the one-to-one mapping interface may diminish over time.

Experience Level

In general, inexperienced participants committed 33% less errors; gave 47% higher usability ratings, and rated tasks to be 46% easier than did the experienced users. In addition, their subjective ratings were higher compared to the experienced participants’ ratings. These results suggest that the reason for experienced participants to give higher ratings is negative transfer of training, since they were already accustomed to browsing the Web. Task completion time for reaching the Web sites was comparable for both experienced and inexperienced users, which suggests that the tree structure was designed independent of previous experience on either computers or the WWW.

CHAPTER 6: CONCLUSION

The main goal of the WWW Browser system is to provide fast WWW access through television for those who either cannot afford the existing hardware or are intimidated by computers. Some of the unique characteristics of this WWW browser system is the use of nine buttons on a universal remote control and an index structure for navigating to Web sites. This research has demonstrated that inexperienced users are more inclined to prefer hand-held internet controllers than experienced ones.

Subjective ratings suggest that users prefer 0.2 second system lag; however, up to 0.7 seconds was considered acceptable. Active feedback was considered the best among the three feedback styles. In addition, it would be advisable to have active feedback in systems such as the Internet since the response times are highly variable.

Overall, the tabbing interface was evaluated to be better than the one-to-one mapping interface. Both the number of errors committed and time taken to accomplish each task were less for the tabbing interface. However, the increase of system lag time had a greater adverse effect on the tabbing interface than on the one-to-one mapping interface. If the indexing structure does not change frequently, it is foreseeable that the significant difference in task completion time and number of errors between the two interface formats will diminish. However, due to the dynamic nature of the Internet, indexing structures based on the WWW are likely to change frequently.

Today, many major electronics companies are marketing products that make the WWW accessible through television and many other companies are expected to follow. Due to the variable response times in the Internet and its dynamic nature, it is essential to consider system lag, feedback style, and interface format when designing such products. This study contains valuable information for systems designers, and may play a role in the future acceptance and popularity of Web access through television.

REFERENCES

- Bergman, H., Brinkman, A. and Koelega, H. S. (1981). System response time and problem solving behavior. In *Proceedings of the Human Factors Society-25th Annual Meeting*, (pp. 749-753). Rochester, NY: Human Factors Society.
- Brown, C. M. (1988). *Human-Computer Interface Design Guidelines*. Norwood, NJ: Ablex Publishing Corporation.
- Dannenbring, G. L. (1984). System response time and user performance. *IEEE Transactions on Systems, Man, and Cybernetics*, **14(3)**, 473-478.
- Eberts, R. E. (1994). *User Interface Design*. Englewood Cliffs, NJ: Prentice-Hall Inc.
- Gallaway, G. R. (1981). Response time to user activities in interactive man/machine computer systems. In *Proceedings of the Human Factors Society-25th Annual Meeting*, (pp. 754-758). Rochester, NY: Human Factors Society.
- Goodman, T. J. and Spence, R. (1978). The effect of computer system response time on interactive Computer-Aided problem solving. In *Proceedings of the 1978 SIGGRAPH Conference*. New York, NY: ACM.
- Goodman, T. J. and Spence, R. (1982). Effects of potentiometer dimensionality, system response time, and time of day on interactive graphical problem solving. *Human Factors*, **24(4)**, 437-456.
- Kiger, John J. (1984). The Depth/Breadth Trade-Off in the Design of Menu-Driven User Interfaces. *International Journal of Man-Machine Studies*, **20**, 201-213
- Kuhmann, W., Boucsein, W., Schaefer, F., and Alexander, J. (1987). Experimental investigation of psychophysiological stress-reactions by different system response times in human-computer interaction. *Ergonomics*, **30(6)**, 933-943.
- Kuhmann, W. (1989). Experimental investigation of stress-inducing properties of system response times. *Ergonomics*, **32(3)**, 271-280.
- Landauer, T. K. and Nachbar, D. W. (1985). Selection from Alphabetic and Numeric Menu Trees Using a Touch Screen: Breadth, Depth, and Width. In *Proceedings of CHI'85*, 73-78.
- Lee, E., Whalen T., McEwen, S. and Latrémouille, S. (1984). Optimizing the design of menu pages for information retrieval. *Ergonomics*, **27(10)**, 1051-1069.
- Lee, E. and MacGregor, J. (1985). Minimizing user search time in menu retrieval systems. *Human Factors*, **27(2)**, 157-162.

- Martin, G. L., Corl, K. G. (1986). System response time effects on user productivity. *Behavior and Information Technology*, **5(1)**, 3-13.
- Martin, J. (1973). *Design of Man-Computer Dialogues*. Englewood Cliffs, NJ: Prentice-Hall Inc.
- Miller, R. B. (1968). Response Time in Man-Computer Conversational Transactions. In *Proceedings of the Spring Joint Computer Conference*, **33**, 267-277.
- Paap, K. R. and Roske-Hofstrand, R. J. (1986). Design of menus. In M. Helander (Ed.), *Handbook of Human-Computer Interactions* (pp. 205-235). Amsterdam, Holland: North Holland.
- Planas, M. A., Treurniet, W. C. (1988). The effects of feedback during delays in simulated teletext reception. *Behavior and Information Technology*, **7(2)**, 183-191
- Schaefer, F. (1990). The effect of system response times on temporal predictability of work flow in human-computer interaction. *Human Performance*, **3(3)**, 173-186.
- Shneiderman, B. (1984). Response time and display rate in human performance with computers. *Computing Surveys*, **16(3)**, 265-285.
- Shneiderman, B. (1988). We can design better user interfaces: a review of human-computer interaction styles. *Ergonomics*, **31(5)**, 699-710.
- Shneiderman, B. (1992). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Reading, MA: Addison-Wesley Publishing Company.
- Snowberry, K., Parkinson, S. R. and Sisson, N. (1983). Computer Display menus. *Ergonomics*, **26(7)**, 699-712.
- Thum, M., Boucsein, W., Kuhmann, W. and Ray, W. J. (1995). Standardized task strain and system response times in human-computer interaction. *Ergonomics*, **38(7)**, 1342-1351.
- Weiss, S. M., Boggs, G., Lehto, M., Shodja, S. and Martin, D. J. (1982). Computer system response time and psychophysiological stress II. In *Proceedings of the Human Factors Society-26th Annual Meeting*, (pp. 698-702). Seattle, WA: Human Factors Society.

APPENDIX A

Prescreening Questions to Determine the Experience Level of Participants

1. Are you familiar with using computers?
 yes no
2. Do you own a computer?
 yes no
3. Do you use computers for (check all that apply)
 personal reasons work/school related reasons
4. What type of applications do you use computers for (check all that apply)?
 word processing (Word, Word Perfect, etc.)
 graphics/drawing (CAD, Corel Draw, Canvas, etc.)
 internet (Eudora, Netscape, etc.)
 programming/ simulation (C, FORTRAN, Sigma, etc.)
 multimedia (Authorware, Director)
 spreadsheets (Excel, Delta Graph, etc.)
 other _____
5. Have you ever used any of the commercial information services such as America On Line and Compuserve?
 yes no
6. Have you ever used the internet to browse the World Wide Web (WWW)?
 yes no
If yes, please continue. If no, please skip to question #10
7. Do you browse the WWW on a regular basis?
 yes no
8. How often do you browse the WWW?
 have never done it 1-3 times a month
 1-3 times a week 4-6 times a week at least once a day
9. Are there sites that you visit on a regular basis?
 yes no
10. Have you ever used the internet for any purpose other than browsing the WWW?
 yes no
If yes, please check all that apply
 e-mail file transfer telnet
 gopher newsgroups other _____

APPENDIX B

Tasks Performed during the Experiment

Practice Tasks

1. Go to the “CNN World News” homepage and read about the latest news from around the world.
2. One of your friends has graduated from the Dimitrie Cantemir High School in Romania. Find the homepage of this school and take a look at the school projects.
3. You would like to order a swimsuit from “Solar Tan Thru Suits”. Go to their homepage and check out the styles for One Piece Suits in the 1996 Collection.
4. Fallingwater is one of the most famous houses of Architect Frank Lloyd Wright. Go to the “Fallingwater” homepage to take a look at the floor plans.
5. You would like to order nuts from “Nuts4u”. Go to their homepage and check out their Gift Tins.
6. Visit the NBC homepage to find out when NBC is showing the TV show “Mad About You” in the fall season.
7. You would like to learn about the religion Baha’i Faith. Go to “The Baha’i Faith” homepage and read about the teachings of Baha’i.
8. Go to “The American Hiking and Backpacking Trail Guide” homepage to learn about the hiking trails in the Zion National Park in Utah.
9. Go to the “BR On Line Travel” homepage and gather touristic information about the city of Sao Paulo which is in the southeast region of Brazil. Brazil is located in South America.
10. Go to the “Brunner Elementary School” homepage and learn about their student projects. Brunner Elementary School is located in New Jersey.
11. Go to the “Institute of Chemical Engineers” homepage to get some information about membership.
12. One of your friends told you about the Bilkent University in Turkey. You are curious about the engineering programs they have at that university. Go to the “Bilkent University” homepage to see what kinds of engineering programs they offer in Faculty of Engineering.

Tasks Performed during Data Collection Stage

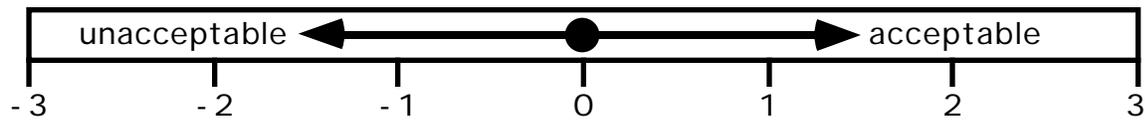
1. You are interested in alternative medicine and would like to get some information about Acupuncture. Go to the “Acupuncture” homepage and learn about the history of Acupuncture.
2. Go to the “Mental Health Net” homepage to learn some tips on keeping your mind “healthy”.
3. Go to the “National Lesbian and Gay Health Association” homepage to read about the health issues related to homosexuality.
4. Kidzeen is a magazine for kids. Go to the “Kidzeen” homepage and take a look at the current issue (issue #7).
5. Go to the “Disneyworld” homepage and find about the type of attractions in the Magic Kingdom Park.
6. You collect stamps and would like to learn more about ships on stamps. Go to “Ships on Stamps” homepage to see the stamp with the sailing warship “Victory”
7. Go to the “Air & Space Magazine” homepage and take a look at the latest issue of the magazine.
8. The harp is one of your favorite musical instruments and you would like to attend the Fourth Annual Harpers’ Escape Weekend in Spring Lake, New Jersey. Go to the “Fourth Annual Harpers’ Escape Weekend” homepage to check out the itinerary for the weekend.
9. You would like to go camping this weekend. Go to the “Camp & Conference Homepage” and learn about the camp of the week.
10. Go to the “Freefall!!!” homepage to find out how you can learn to skydive.
11. Go to "The Official Dennis Rodman Homepage" to take a look at the Rodman Hair Archive. Dennis Rodman is a basketball player.
12. Go to the Embassy Suites homepage. Embassy Suites is the largest all-suite, upscale hotel chain in the country.
13. You would like to go for a Safari trip in Zambia, Africa. Go to the “Travel & Tourism in Zambia” homepage to learn about the accommodations in Lower Zambezi National Park.
14. Go to the “Swissair” homepage to learn their schedule for flights to the US.
15. Go to the “Encyclopedia Britannica” homepage to read about the history of Egypt.
16. You need to get financial aid for college education. You have learned that the company “Hunter and Associates” may have some money available for you. Go to their homepage to get some detailed information about eligibility requirements.
17. Your nephew is gifted with high IQ. Therefore, you want to find out about Education Program for Gifted Youth. Visit their homepage and learn about the courses they offer.

18. Find out information about the McIntosh High school in Georgia.
19. Go to the “Boeing” homepage to find information about its new commercial aircraft 777.
20. Go to the “FORTRAN” homepage to learn about the latest updates. FORTRAN is a programming Language.
21. Go to the “Virginia Tech” homepage and find information about Graduate School admissions.
22. Find the “White House” homepage in the Executive Branch of the Government and take a tour of the President’s home.
23. Go to the “American Citizens Abroad (ACA)” homepage and learn about overseas voter registration.
24. Find the “Hawaii State Government” homepage and meet the Governor.
25. Go to the “Career Expo Job Fairs” homepage and get information on the upcoming job fairs.
26. Go to the “New York Stock Exchange” homepage and learn about the companies that just joined the ranks of New York Stock Exchanged-listed stocks.
27. Go to the “United States Air Force” homepage to take a tour of the Air Base.
28. You are doing your final shopping before you go on a trip around the world. Go to the “World Traveler” homepage and check out their travel accessories.
29. You have heard about the natural skin care products of “A Fashion Experience”. Go to their homepage and find out how you can order their products.
30. You want to buy some aquarium supplies. Go to the “Aquarium Technology” homepage and learn about their November Super Specials.

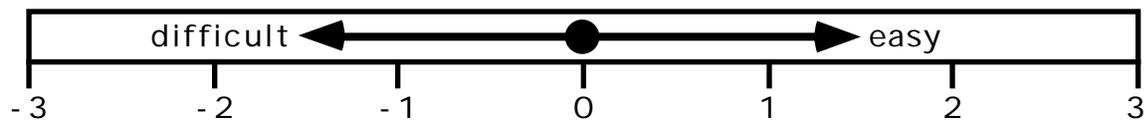
APPENDIX C

Subjective Evaluation Questions

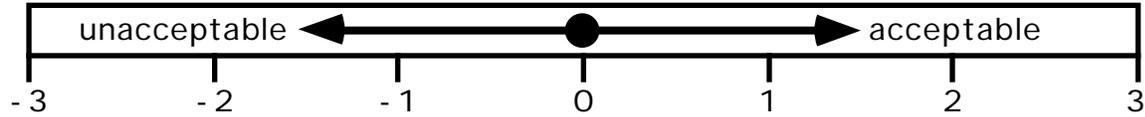
1. Responsiveness of the system (time it took for the system to respond)



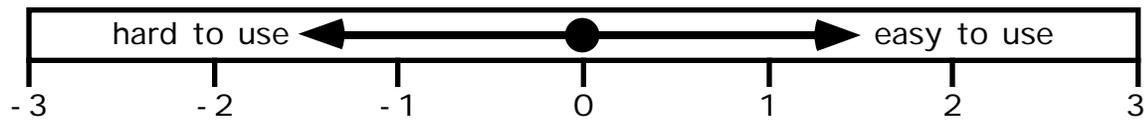
2. Difficulty of completing the task



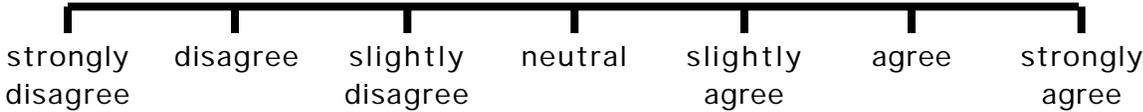
3. Acceptability of the information that the feedback provided



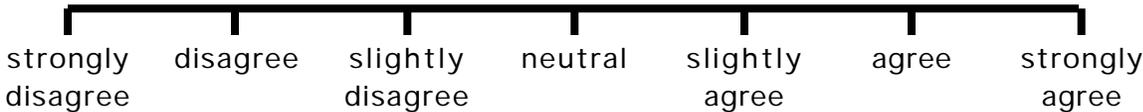
4. Usability of the interface design



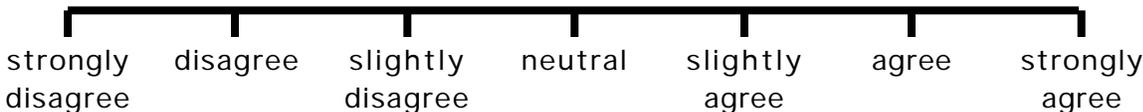
1. The text on the screen was legible.



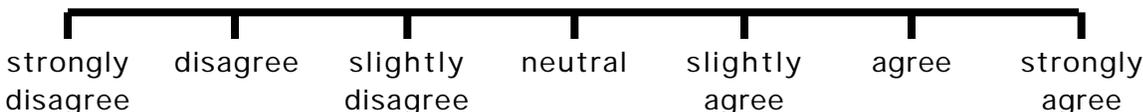
2. The organization of the topics in the tree structure was easy to follow.



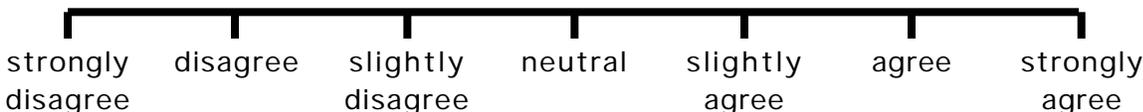
3. The number of topic choices on the screen was acceptable.



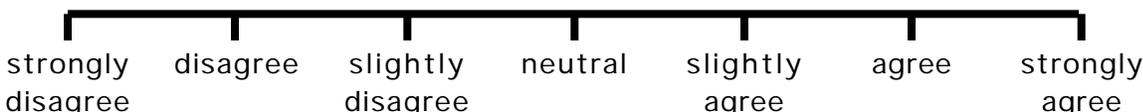
4. It was easy to navigate through the tree structure by using the nine buttons.



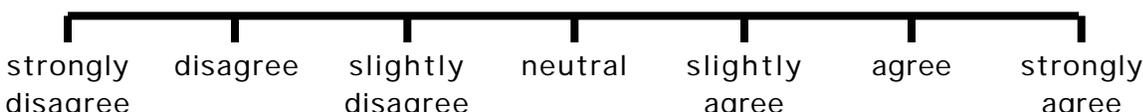
5. It was easy to adapt to the change of button functions from nine buttons on the screen to nine buttons with assigned functions (applies only to the One-to-One Mapping Interface).



6. The button legend on the screen was easy to understand.



7. The button legend on the screen was helpful.



1. I prefer

no feedback

passive feedback (man running in place)

active feedback (man running, bar changing color, percentage changing)

2. I prefer

the One-to-One Mapping Interface (the one with nine corresponding buttons on the screen where the ninth button is always a “Cancel” button).

the Tabbing Interface (the one that has assigned functions for each button on the remote control and requires one to tab through the topic choices on the screen)

VITA

Feyza F. Ergen was born in Istanbul, Turkey on February 17th, 1973. She obtained her Bachelor of Science in Industrial and Systems Engineering from Virginia Polytechnic Institute and State University in May 1995. During graduate school, she was a teaching assistant for Work Measurement and Methods Engineering course and a research assistant in the Displays and Controls Laboratory. Upon completion of her Master Science, she will be employed as a Product Design Engineer by Ford Motor Company in Dearborn, Michigan.