Design and Testing of an Adult Age-Independent Online Needs Assessment Tool and Development of Adult Age-Independent Design Guidelines

Tolga Durak

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

Tonya L. Smith-Jackson, Ph.D., Chair
Maury A. Nussbaum, Ph.D.
Karen A. Roberto, Ph.D.

April 24, 2003
Blacksburg, Virginia

Keywords: Adult Age-Independent, Online Needs Assessment,
Design Guidelines, Prototype Heuristic Evaluation

Design and Testing of an Adult Age-Independent Online Needs Assessment Tool and Development of Adult Age-Independent Design Guidelines

Tolga Durak

(Abstract)

This thesis addressed the following research activities: (1) Developing and evaluating the usability of an adult age-independent online assessment tool that includes health care related content and is centered on the needs and capabilities of both older and younger users and, (2) Developing adult age-independent design guidelines. An online tool, which was developed for a previous study, was revised according to the results of the previous study and preliminary development activities based upon the user requirements of older adults (50 years or older). The online assessment tool was also tested with younger adults (between 18 and 35 years of age) to determine the differences between younger and older adults. Subsequent to these activities, a final tool that captured the needs of both older and younger adults was developed and evaluated for usability. According to the usability testing results, a design guideline set for an adult age-independent Online Needs Assessment Tool was developed and revisions were conducted in order to develop the final Online Needs Assessment Tool.

The results of these studies, along with the tools developed, provide online survey and interface designers information on older and younger adult user requirements. They provide a knowledge resource for older and younger user healthcare information needs and information about the usability of one online assessment tool to accommodate both older and younger adult users. They also assist interface designers to proliferate and to facilitate more rapid application development.

Table of Contents

Table of Figures

Table of Tables

Dedication

This paper is dedicated to my grandfather, Mehmet Saim Yilmaz (as known as "Saim Hoca"), who would have been proud.

Acknowledgements

This thesis would have not been possible without the help of my advisor, Dr. Smith-Jackson. She did everything to help me while I was writing my proposal, running my study, and completing my thesis. If it were not for her vast Human Factors Engineering knowledge and experience, positive support and warm welcome, this project would not be completed.

I owe so much to my parents, who have always encouraged me in my academic life, helped me with every decision, and made being a good student enjoyable. They did nothing but their best to ensure the best education and the best career choices for me. Without their never-ending love and support, I would not have made it this far. To my parents, who shared their love, respect and admiration for engineering and hard work.

Thanks to my committee for always having positive, helpful feedback, for always having confidence in my abilities, and for being flexible in rescheduling timelines to accommodate technical and personal difficulties. Thanks to Miranda Capra and Curtis Jackson, for helping me with server and coding problems. Thanks to Erdem Agan for being the brother I never had. Thanks to Chris Allen Chacey for showing me the true meaning of “phi phi kappa alpha” in these very stressful times.

1. Introduction

The Internet is a tool that serves a wide range of purposes for different users. Internet applications not only allow users to perform several activities, but they also have an effect on users' psychosocial well-being. In a recent study, the electronic mail access system, which is one of the common Internet applications, was found to improve psychosocial well-being among older users by decreasing their loneliness (White, McConnell, Clipp, Bynum, Teague, Navas, Craven & Halbrecht, 1999).

One of the main uses of the World Wide Web (W W W) among younger and older adults is health information access (Morrell, 2000). According to Morrell, while using the W W W, younger and older adults primarily prefer to learn how to use electronic email and seek information about health topics and traveling. The preference of older users to seek health information while using the W W W, seems to be the main motive behind new online applications in the fields of needs assessment tools and medical support. The development of new online applications in these fields (needs assessment tools and medical support) will increase the breadth of Internet usage and broaden the range of the user population. The development of these applications will also accommodate the needs and interests of older user population, which has been found to account for 17% of the Internet users in a recent study conducted by Microsoft and the American Society on Aging (Associated Press, 1999).

Despite the contextual similarities needs assessment and usability have, they show differences in functionality and areas of use. Usability can be defined as effectiveness, efficacy and satisfaction necessary for a specific user group to accomplish a certain set of tasks (Garner, 2000). Needs assessment on the other hand

is evaluating the needs of a certain user group to optimize a current system (Reviere, Berkowitz, Carter & Ferguson, 1996). Needs Assessment Tools are devices that support systematic exploration of the current system and optimum system to be developed (Reviere, Berkowitz, Carter & Ferguson, 1996). Online Needs Assessment Tools (ONATs) are a form of needs assessment tools, which are distributed by using a network system. The development of these online tools seems to be increasing in frequency in the field of medical support. A statistical analysis was performed by the U.S. Census Bureau that yielded a list of the most researched Internet topics by the older adult population (50 years or older). The U.S. Census data revealed an increase in the frequency of using the W W W for medical purposes. According to this representation, in 1993, health care issues were not even listed (U.S. Census Bureau, 1993). Therefore, the topic was not even searched enough to make the final statistical readings. In 1997, however, older users ranked third in terms of percentage of healthcare information that was researched on the Internet with 76.5% compared to 82.1% for 25-49 year olds and 78.5% for 18-24 year olds (U.S. Census Bureau, 1997). Health care information was also in the top four topics researched on the Internet for all age groups. The other three topics were email, government issues, and business issues.

Health on the Internet (Hon) Foundation also published several research findings that emphasized the preference to access online healthcare information. Based on the results of several surveys on the use of the Internet for medical and health purposes among 4,368 participants used, Hon Foundation (1999) stated three major trends:

1) Younger adults are the largest group that seeks medical and healthcare information on the Internet; older adults comprise a significant amount of the user population.
2) There is an increasing amount of users who use of the Internet for medical and healthcare purposes everyday.
3) A large percentage of Internet users think that the quality of medical and healthcare information should be improved.

The increasing interest in health information on web sites, the need expressed for quality improvement of online information, and the usefulness and efficiency that online applications provide have led ONATs to receive more research attention and to be used more frequently. In addition to the common characteristics all needs assessment tools have, ONATs are capable of providing users with easy access at any time of the day or any day of the week. The service is time-efficient compared to face-to-face methods and can reach potential participants in isolated environments or those who lack of transportation. The tools are eliminating the problems that personal direct medical assistance might have, such as restricted mobility caused by health issues or fear of crime (Smith-Jackson & Williges, 2000). In addition to the common advantages that all online applications have, these tools can support the acquisition of user-centered guidelines for design.

Due to the preference to access healthcare information on the Internet compared to any other type of information, the increase in the frequency of using the W W W for healthcare and medical purposes, the advantages web-based applications have over traditional applications, and the need expressed for quality improvement of healthcare

information on web sites, there is a need for online assessment tools designed for older and younger adults. There is an overwhelming amount of Web page applications that are health-related for older users. The W W W offers web sites related to older users (e.g., *seniorresource, healthfinder-just for you seniors, seniornavigator*), or sites related to health or aging (e.g., *webmd health, drkoop, carescout or national institute on aging*). The literature consists of several studies and empirical results that support the basis of features necessary for these web site applications. However, there are only a few online assessment tools to conduct formal needs assessment to develop these applications. In addition, none of these assessment tools are developed for older users or are adult age-independent.

The only type of ONAT that has received research attention is the online survey. There are many online survey applications developed by institutes and companies to determine user demographics or satisfaction. Feinberg and Johnson (1997) have conducted a literature review for their customer satisfaction survey project. The surveys were reviewed under the categories of demographic surveys and user satisfaction surveys. Under the first category, five user surveys of the Graphics, Visualization, and Usability (GVU) Center of Georgia Institute of Technology, the CommerceNet/Nielsen Internet Demographics survey and the American Internet User were reviewed. Under the category of user satisfaction, customer satisfaction surveys of the National Technical Information Service, Usajobs, ERIC Document Reproduction Service, and University of Illinois Customer Satisfaction Survey were reviewed. Amongst the surveys reviewed, only the GVU Center of Georgia Institute of Technology provided empirically-supported guidelines. These guidelines were for developing Web based surveys that

have a goal to determine demographic information of W W W users or customer satisfaction with a product or service (see Appendix A for guidelines). By establishing guidelines based on their research results, GVU Center of Georgia Institute of Technology tried to decrease researcher and designer biases. These user-centered guidelines will not only help prepare new surveys, but will also provide guidance for further research on determining the efficacy of online surveys used to capture user needs. GVU Center of Georgia Institute of Technology also demonstrated the use of online needs assessment to identify design guidelines.

Since the focus of this research includes online assessment for older users, the research domain of older users and technology is applicable. In the research domain of older users and technology, studies have been conducted to characterize and define the special considerations of the older population and technology usage. Some of the studies were performed to determine the changes in cognitive, perceptual, and spatial ability that older users experience. The results of these studies will not be discussed in greater detail since they are not specifically pertinent to the scope of this discussion.

The rest of the studies in this research domain focused on the effects of the lack of computer experience and on the stress caused by computer tasks on the older population. The studies have shown that the general lack of experience with computers can cause older users to develop biases towards new technology applications. A study by Danowski and Saks (1980) found that older people tend to have positive attitudes towards new technology if they had positive experiences in the past. In addition, Brickfield (1984) found that elderly users were less willing to use technological products if they did not have positive attitudes towards them.

Besides the biases that the older user might have developed towards new technology applications, new technological applications in the form of online tools (to be specific) might cause intimidation. If developers do not take into account the amount of additional stress that might be induced by a computer task for older users compared to younger users, lack of computer confidence and intimidation for the older user could occur (Czaja, 1997). This lack of computer confidence or intimidation could potentially be the biggest contribution to older user apathy.

Older users, their interactions with computers in general and the factors mentioned above (i.e., changes in cognitive, perceptual and spatial ability, effects of computer experience and stress caused by computer tasks, etc.), need to be considered while designing for older users. However, the degree of access to and adoption of computers by older adults could be one of the most critical factors to consider when using online tools. One justification for this reasoning relates to the pace of technology change. A look at technology development in the past ten years reveals that technology has evolved from a Commodore 64 computer to a Mac G4 (Moravec, 1999). Even the concept of the Internet was a vague definition 8-10 years ago. When considering these rapid changes in the technology domain, access to and use of the new products or inventions can be extremely time and resource dependent, especially for the older population.

As previously discussed, there is a need for ONATs to improve healthcare information delivered via the Internet for all age groups. Even though few studies have been designed to accommodate older user needs, none of these studies covered all or most of the aspects and the requirements that are needed in order to develop the whole

ONAT for older users or for users of all adult age groups. In order to contribute to the research in the fields of medical and online applications, this study developed and evaluated the usability of an adult age-independent ONAT that includes healthcare related content. The development and evaluation of Adult Age-Independent ONATs met the need for ONATs for all adult age groups. The use of adult age-independent tools accommodated the user requirement differences between older and younger users. In addition, the use of these tools is expected to decrease the time and resources needed during the development stage and to provide design requirements that can be implemented in similar currently-used applications in order to accommodate both older and younger users with the same application. The subsequent literature review highlights relevant studies in the field of gerontology and gerontechnology that relate to user-centered design for older compared to younger adult users.

2. Literature Review

User-centered design activities do not typically involve older users. Consequently, there is a lack of applications designed for older users. There are three possible reasons behind the lack of design applications targeted for older users. The first possible reason is the prevalence of use among the younger population, who are often seen as the potential customers in the technology industry. The applications are usually based upon user requirements of the potential consumers, most of whom are believed to be younger adults. The second possible reason is that the majority of design teams are composed of young adults. The younger designers will have an impact on all steps of the design process involving the development of a conceptual model.

In every design application, the designer must develop a conceptual model that captures the significant aspects of the application (Norman, 1990). Developing the conceptual model is an iterative design process where the designer has to communicate the user's mental model into the system image and its functionality. Since the designers and programmers in the industry may not typically have time or resources to conduct research during the conceptual model development phase, they rely on past experiences and previous research to capture the user's mental model for the particular system. According to Norman (1990), there are major differences in formulation of experience, perception, cognition and benefits of different training instruction methods (basically the difference in mental models) between older and younger adults in general. If necessary research is not conducted to capture the user's mental model for a system, these differences make it difficult to capture the system image that is understandable by older users (Norman, 1990). As a result, the development of a conceptual model and

the effect of mental model differences cause the applications to be unsuitable to older users' needs. These incompatible applications might cause older users to get confused, to require more adaptation time, or to be forced to develop new mental models that are not intuitive.

The third possible reason is the misconception about the increasing sales revenue of older users as customers in the fields of computing and Internet technology. Several studies provide evidence to support that older users account for a significant percent of the potential customer population. In 1998, Nielsen Media Research and CommerceNet stated that older users (who were defined as users 50 years or older) were the second fastest-growing Internet user group after 16-to-24 year-old user groups (Associated Press, 1999). Another study by Microsoft and the American Society on Aging revealed that older users (50 years or older) formed 17% of the Internet users (Associated Press, 1999). Although older users account for a significant percent of the customer population and the percentile is constantly growing, few design activities involve older users (or their needs and requirements) since they are not perceived as the potential customers. However, older users are recognized to be a special user population by various researchers, and thus, are treated as a marginalized group. Several research investigations have been conducted to identify special design requirements for older users in the field of gerontology and gerontechnology. Also many companies (such as America Online and Microsoft) have begun to view older users as one of the final unused market resources (Associated Press, 1999). These companies have begun to routinely conduct studies in order to build products to accommodate older users.

Research that has been conducted related to older users can be classified under two categories. The first category of studies consists of research conducted to determine the physical and psychological capabilities of older users such as vision, hearing, hand function, and changes in mental activities (e.g., Haigh, 1993; Hawtorn, 2000). The second category consists of research conducted to test or to evaluate design applications such as automatic teller machines, word processing interfaces, email or Web browsing (e.g., Adams & Thieben 1991; Charness, Schuman & Boritz 1992; Czaja, Clark, Weber & Nachbar, 1990; Ellis & Kurniawan, 2000; Rogers, Gilbert & Cabrera, 1994).

In the following section the implications of research results and their relevance to the ONATs in the fields of gerontology, gerontechnology, and their sub-categories will be discussed in greater detail.

2.1 Gerontology and Gerontechnology

Gerontology is the scientific discipline that studies the biological, psychological and sociological phenomena associated with old age and aging (The American Heritage College Dictionary, 1993). In contrast, gerontechnology is the scientific discipline that uses applications of technology to improve the living and working environment and medical care of aging and aged people (Brouwer-Janse, Suri, Yawitz, Vries, Fozard & Coleman, 1997). In the field of gerontechnology, designers and architects try to find ways to adapt to the challenges and opportunities of aging. That is, gerontology is the primary field that acquires knowledge of the special needs of older adults; gerontechnology is the field that applies information gained from gerontology to the design of technology and systems.

According to Brouwer-Janse et al. (1997), consumer involvement (user-centered design) is defined as the most important aspect in technology development and adaptation of technology to the lives of elderly users. Also Brouwer-Janse et al. (1997) stated the main goals of gerontechnology as follows:

- Aging research improvement: More research should be conducted in the field of aging on specific aspects as well as the general concepts and applications.
- Prevention: Technology should serve to prevent the problems that older users face.
- Enhancement: Technology should enhance the lifestyle and the standards of the older users.
- Compensation: Technology should compensate for the decrease in older user abilities.
- Care giver assistance: The care giving equipment should be designed to be suitable, flexible and easy to use.

In these scientific disciplines, studies relevant to the proposed research, which fall in the fields of Gerontology and Human Factors, are concentrated in the following sub-categories: 1) Aging and Capabilities, 2) Design Applications for Older Users, and 3) Younger Users versus Older Users.

2.1.1 Aging and Capabilities

When designing the interface of the ONAT, several factors need to be taken into consideration. These items are: text size, text spacing, background color, use of graphics, and visual field. When designing for older users, the findings from studies suggested several adaptations. Some of these findings have suggested that text size

should be set in the 12 – 14 point size range, background color should be void of the blue/green range, graphics should be kept to a minimum or voided altogether, and the visual field should be placed in narrow view (Hawthorn, 2000). In addition, distorted spacing in the text was found to cause a distraction that led to an undesired emphasis on particular words or phrases (Shaw, 1990). To avoid this distraction, the block justification style, which automatically provides equal margins from the beginning and the end of the sentences by editing the spacing between the words, should not be used. Another prominent factor that creates new considerations for the interface design is the change in the visual short-term memory, which reduces the ability to utilize a perceptual organization accurately while performing a recall task (Humphrey, Kramer & Gore, 1994). Humphrey et al. (1994) found that older adults have a poorer ability to utilize grouping factors such as incomplete, embedded or partial figures, charts, tables or reports in visual design applications.

2.1.2 Design Applications for Older Users

Design applications for older users can be classified in the following categories: Older users and ATMs, Interface Design Applications for Older Users and Other computer related applications for older users.

2.1.2.1 Older Users and ATMs

In one of the first studies on older users and automatic teller machine interaction, Adams and Thieben (1991) found that transaction demonstration, which was used with an active task, on the ATM simulator without any hands-on experience, had significant gains for elder users. The results of this study were extended by other studies concentrating on different aspects of automatic teller machines and older user

interaction. Rogers, Gilbert & Cabrera (1994) conducted an in-depth structured interview with elderly users on automatic teller machine usage. The results of these interviews revealed the frequently performed tasks and the issues that caused major difficulties. The frequently performed tasks were withdrawals, deposits and balance checks. Although the tasks that older users performed frequently are not pertinent to ONATs, the difficulties that older users encountered contribute to and highlight the importance of the research in the field of Online Needs Assessment. In their study, Rogers, Gilbert & Cabrera (1994) found that regardless of the experience level, older users had difficulties with inserting the card, recalling the pin number, recalling the steps of the tasks and adapting to different ATM interfaces. Older users also expressed their concerns for safety (such as ATM machines giving the wrong amount of cash, not giving any cash at all, or fear of crime during money transactions outside since most of the ATM machines are located outside) and feeling intimidated by making other users wait in line. The results of this study highlighted the importance and necessity of all online tools. The online tools will provide flexibility so that the older users will be able to perform necessary tasks at their convenience and set or change the user preferences. The tools will also eliminate any time constraints and provide one-on-one interaction that will avoid the intimidation of older users. In addition to all these advantages that all online tools have, ONATs also could be used to gather information on the needs, capabilities and preferences of older users to eliminate the cognitive and physical difficulties users might have.

Rogers, Fisk, Mead, Walker and Cabrera (1996) have also studied the automatic teller machine usage by using different training methods (i.e., various forms of

instructions or walk-troughs, practice sessions). Even though the findings were case specific to ATMs, a general requirement can be constructed for other interface applications designed for older users. The results of these studies support the general requirement that design applications for older users that incorporate help-aid tools (in the form of instructions, walk-throughs, practice sessions or demonstrations) increased the performance of older users.

2.1.2.2 Interface Design Applications for Older Users

There has been only a small amount of research conducted in the field of interface design applications for the older users. These studies have usually concentrated on a particular aspect of the interface design. In this area, Worden, Walker, Bharat and Hudson (1997) conducted research on two different interaction techniques. The techniques were using area cursors, which are “cursors that have a larger than normal activation area.” (p. 267), and sticky icons, which are “icons designed to have an automatic reduction of the cursor’s gain ratio when the cursor is on the target icon.” (p.267). It was also suggested that the decrease in cursor’s gain ratio would increase the effective target size, making it easier to place the cursor on the target. In their study, Worden et al. (1997) tested the effects of these two different techniques (using area cursors and sticky icons) on performance enhancement individually as well as combined together. The results showed that the combination of these two techniques (using area cursors and sticky icons) decreased target selection times for older adults by 50% compared to using each technique individually. These results, in relation to the interface design, suggest that the use of area cursors and sticky icons at the same time should be integrated in every interface design developed for older users for

performance enhancement. Consistent with the results of Worden et al.'s (1997) study, for almost every software interface, area cursors and sticky icons are used at the same time.

Nielsen and Schaefer have conducted a study on the auditory aspect of interface design. As a result of their study, Nielsen and Schaefer (1993) found that the addition of audio to the interface had different impacts on older and younger adults. The addition of audio seemed to be enjoyable for younger users, while older users were overwhelmed by these effects and found the tasks more complicated. The results of this study were also supported by the findings of McDowd and Birren's study (1990), where the additional auditory or visual cues (in the form of tones) were found to cause a time delay and interruption of concentration of the older users for the current task compared to younger users. These results, in relation to the interface design, suggest that no auditory or visual cues should be used in design applications developed for older users in order to avoid time delay and interruption of concentration.

2.1.2.3 Other Computer Related Applications for Older Users

Many studies have been conducted in the area of the general uses of computers and Internet. The results gathered from these studies provided valuable information for new design applications as well as further research in the field of ONATs.

In their study, Czaja, Clark, Weber and Nachbar (1990) examined computer communication among older adults by installing a customized email system in the homes of 38 elderly women users. The results of this study showed that older users are willing to use the new technological systems if the usage itself is easy, simple, and the features are added over a period of time rather than all at once, and a help support is

provided. Ellis and Kurniawan (2000) also conducted a study in the field of computer communication. By their participatory design, Ellis and Kurniawan (2000) tried to reconstruct an existing World Wide Web site to make it more user friendly for older users with format and display modifications. From this study, prototypes were developed and tested, which supported the development of some general design guidelines (see Appendix B for guidelines).

Beyond the studies in the fields mentioned above, researchers have also developed design guidelines and suggestions to target the elderly population from the findings of these empirical studies. Brouwer-Janse et al. (1997) have had suggestions on applications to help the older population become involved in the communities in which they live and take full advantage of new technology products. These suggestions were generalized improvements in versatility (improvements in the independence of the application usage from any form of help), autonomy (improvements in automation), and social involvement (improvements in the amount of social activities and interaction that involve the user).

2.1.3 Younger Users versus Older Users

The majority of design applications are aimed towards targeting the younger user population. It would be beneficial to distinguish the design commonalities and differences between older and younger user populations when discussing interface designs. This will ensure the modification of current applications in use to capture the needs of the older populations. It will also allow the development of adult age-independent design implementations. Adult age-independent design implementations will not only accommodate the needs and requirements of different adult age groups,

but it will also provide efficiency by cost savings due to the decrease in the number of different development cycles and tests for each adult age group.

In addition to the changes in physical capabilities, several researchers have tried to gain more understanding of changes in cognition speed and accuracy that older users encounter. Vercruyssen, Carlton and Diggles-Buckles (1989) have found that the response selection stage of information processing differed for older users compared to younger users. As a consequence, it took twice as much time on average for older users to respond to a stimulus compared to younger users. Vercruyssen et al. (1989) also showed that the older subjects required more time to derive or obtain the necessary information from the information source. In this study, reaction time was examined for different factors such as age, gender, activation, stimulus degradation, and practice on attention and visual choice. It was also stated that the response time might not be able to reflect the cognitive process, since it was not possible to determine the effects of interactions between multiple factors.

In the area of working memory, Cooper, Lee, Goska, Anderson, Gay, Fickes & Fisk (1992) found that performance with the older adult population was improved with consistent use of similar training programs. This, in relation to the interface designs, suggests that new Web application software and training programs should be designed to be consistent with computer application software programs that are currently used so that consistency could ensure improvement in the older adult population performance.

Charness, Schuman and Boritz (1992) also have conducted studies in the field of computer training for older users. These researchers examined the effects of training technique, computer anxiety and age while training older and younger adults in word

processing. The results of this study not only showed that younger users demonstrated a higher performance level than older users, but it also indicated that self-paced training led to higher performance outcomes than fixed-paced training for both groups of users. Older users required more time to complete the same set of tasks compared to younger users even when computer experience and typing speed were controlled (Sharit & Czaja, 1994). The self-paced training eliminated the time delay. Also, the tutorial, which was incorporated in the self-paced training, provided help and assistance. The tutorial provided a hint when no key was pressed in 30 seconds or an erroneous key was pressed. These studies expressed the importance of constructing the design of the ONATs with tasks that will not be limited by time constraints and will incorporate help-aids for older users. In addition, Zandri and Charness (1989) found that during software training, although older users achieved almost the same performance level, they took twice the time and requested help two or three times as frequently compared to younger users. Even though interface designs of online tools are simpler, to the point, and not timed simulations, the results of this particular study (Zandri & Charness, 1989) suggests an additional requirement to be considered during the development of online tools. The results support the general requirement that design applications for older users should only be designed such that a single, predominant task is present but also peripheral tasks, if any, are easy to do.

The suggested requirement that applications for older users should be designed to have a single task at a time was also supported by results of Sit and Fisk's study (1997) in the area of performance differences. In their study, Sit and Fisk (1997) showed that while young adults were able to allocate sufficient attention to all multiple

task components, older users' performance decreased due to task complexity. Also Sit and Fisk (1997) suggested that this outcome was due to divided attention and the loss of ability to effectively allocate attention. In order to prevent the occurrence of additional mental workload caused by divided attention, any application or design should incorporate a single task at a time.

2.2 Summary of Literature Review – Research Purpose

According to the preference to access health information on the Internet compared to any other type of information, the increase in the frequency of using the W W W for medical purposes, the advantages web-based applications have over traditional applications (as mentioned in previous discussions), and the need expressed (by adult users of all age groups) for quality improvement of healthcare information on web sites, it was seen that there was a need for ONATs that are:

1) Built by applying user-centered design principles for users of all adult age groups.
2) Used to determine the current inadequacy of online healthcare information delivered to users of all adult age groups.

An age-independent ONAT was the only tool that would meet both of the needs stated above. In addition to meeting both of the needs, an age-independent ONAT would also provide a more composite structure. This structure would be more time and resource efficient in the design development cycle and would benefit all age groups. The structure would also distinguish the healthcare information preferences of older and younger adult users and would accommodate the differences in user requirements for both younger and older adult users.

During the development of the ONAT, existing user-centered design guidelines (all guidelines that apply to ONATs) were implemented to capture the user requirements of older adult users while a separate test was be conducted to determine the user requirements of younger adult users. In addition to the goals stated above, this project

also tested the usability of the adult age-independent ONAT in order to make necessary modifications and improvements.

In the first section of Study 1, a literature review was extended in order to gather all necessary user-centered design guidelines. Implementation of these design guidelines was used to capture the user requirements of older adult users. The online needs assessment prototype (Prototype A), which was developed for a prior study, was revised by applying the results and findings of previous research of Smith-Jackson Williges, Kwahk, Capra and Durak (2001) (see Appendix C for the summary of results) and other user-centered design guidelines gathered from the literature review.

Subsequent to these preliminary development activities, the second section of Study 1 was conducted in order to develop design guidelines to capture the user requirements of younger adult users. During the second section of Study 1, the online needs assessment prototype, which was developed during the preliminary development activities, was tested with younger adult users. According to the results and the findings of this usability testing, no new design guidelines for younger users were generated. However, several changes were made while conducting the first design iteration in order to accommodate the user requirements of younger adult users.

The usability of the improved prototype was tested with younger and older adults in Study 2. The results of the usability testing were used to develop usability guidelines for an age-independent ONAT. The results of usability testing were also used to conduct the second design iteration. The second design iteration revised the existing ONAT to better support the user requirements of older and younger adult users.

All design activities proposed in Study 1 & 2 were the formative evaluation stage of a Final Online Need Assessment Tool. Study 1 & 2 consisted of the design and development cycles of the ONAT before the release of the product for use, which is defined as formative evaluation by Preece (1993). During the formative evaluation stage, two design iterations were made to the assessment tool. The reason for selecting two design iterations was due to the finding that testing three different user-interface versions of a product is found to detect most of the usability problems during the formative evaluation stage (Nielsen, 1993b). Since Prototype A, which was one user interface version of the ONAT tested in a prior study by Smith-Jackson et al. (2001), only two other user interface versions of the ONATs were tested in the proposed studies (Prototype B and Revised Online Needs Assessment Tool 1).

Usually in product design and development, the summative evaluation stage comes after formative evaluation. However, no summative evaluation activities were proposed in this research because of time and resource limitations for long-term iterative design activities. Further research could focus on release of the tool and subsequent summative evaluation.

2.3 Research Objective

The primary research objective of this study was to develop usability guidelines for online applications with healthcare information content. Even though several researchers have identified usability guidelines for certain age groups, none of the studies have examined the development of guidelines that would accommodate users of all adult age groups. Therefore, during this study the usability guidelines that are applicable for online applications with medical content were gathered by using an extensive literature review and usability testing. These guidelines were used to support the needs, requirements, and expectations of the older adult user population and younger adult user population separately as well as collectively. These guidelines were also implemented to iterate several ONAT prototypes and tested for accuracy. As a result of these research activities, the final set of usability guidelines that accommodated the user requirements of older and younger adults was developed.

The secondary research objective of this study was to develop a usable Age-Independent ONAT. The only ONAT that have received research attention are the online surveys. There are many online surveys in current use. Some of these surveys are based on empirical findings or suggested guidelines. Some researchers have also provided design suggestions and guidelines for certain age groups, but no studies have examined the development and usability evaluation of adult age-independent needs assessment tools. Therefore, this study was also used to determine: To what degree can one ONAT be developed to be usable for both older and younger adult users.

3. Methods

Two different studies were conducted. The first study was composed of two different sections. The first section consisted of gathering and implementing the findings of other user-centered design research for older users (preliminary development activities). At the end of the first section Prototype A (The existing ONAT) was revised to develop Prototype B. The second section of Study 1 consisted of conducting usability testing and extracting guidelines. In the second section, Prototype B was tested with younger adult users to capture their user requirements. According to the results of this usability testing, guideline extraction criteria were used in order to extract new guidelines and form Design Guideline Evaluation 1 (see Section 3.1.4 for more details).

For study 2, according to Design Guideline Evaluation 1, the first design iteration was made to structure the Revised Online Needs Assessment Tool 1, which was used to support the user requirements of both older and younger adult users. The Revised Online Needs Assessment Tool 1 was also tested with older and younger users to determine the usability of the tool. Subsequent to the usability testing, guidelines extraction (see Section 3.1.4 for more details) and the second design iteration was conducted to improve usability of the Revised Online Needs Assessment Tool 1. After the second design iteration, Revised Online Needs Assessment Tool 1 was revised to develop Revised Online Needs Assessment Tool 2. The overall flow of events for both of the studies is illustrated in figure 1 on the next page:

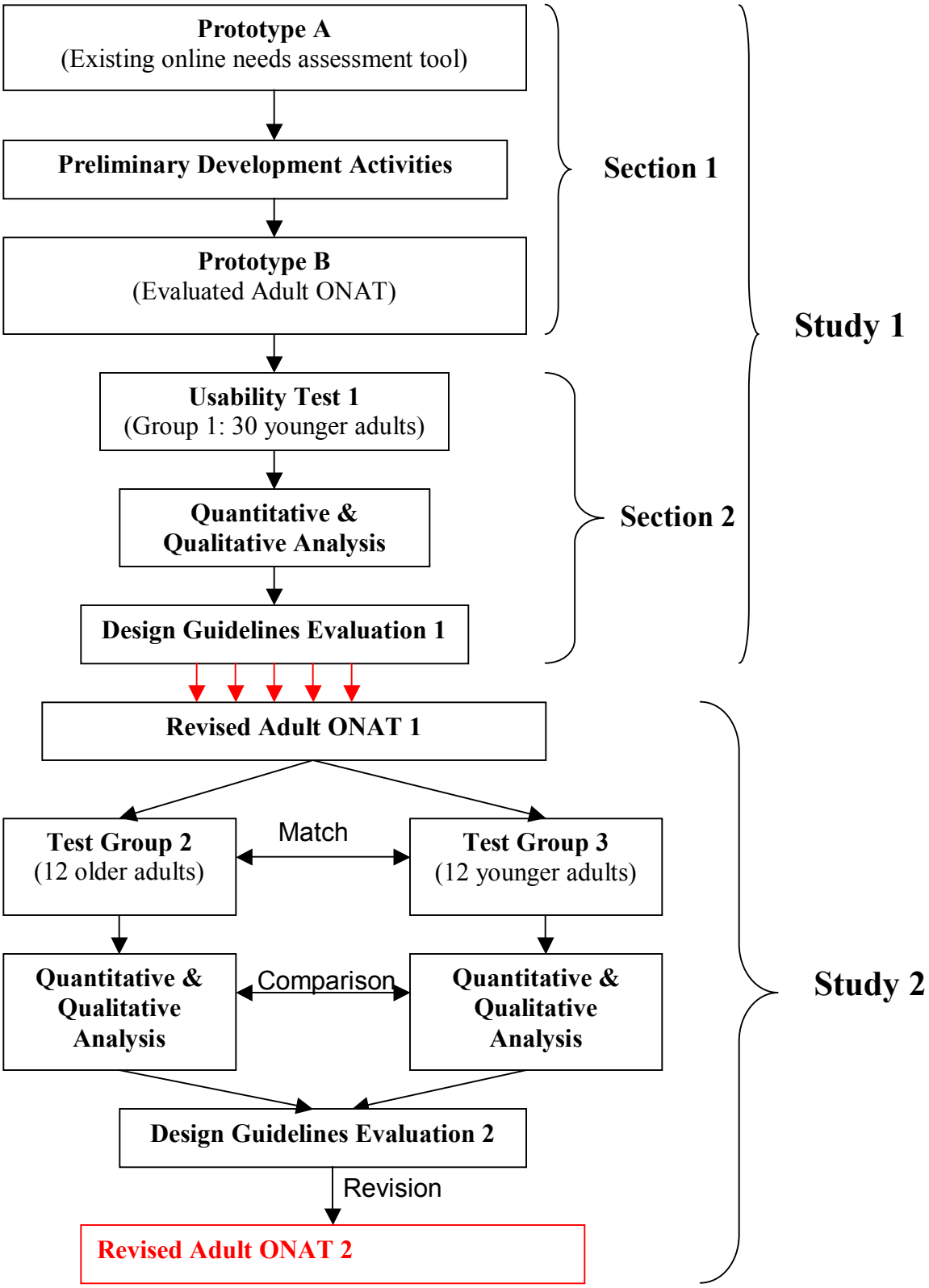


Figure 1: The overall flow of studies.

3.1 Study 1

As mentioned in the prior discussions, Study 1 consisted of preliminary development activities (Section 1), usability testing and extracting guidelines to capture the user requirements for younger adult users (Section 2).

3.1.1 Participants

Thirty participants were selected from the younger adult user population (between 18 and 35 years of age) from the local community. All participants were required to have no major hearing or vision loss that interferes with daily functioning. All of the volunteer participants were be given the participant categorization tool (see Appendix D). With the help of the categorization tool, the potential participants were categorized into novice, intermediate, and expert users. The additional selection and assignment criteria for the participants will be discussed in the research design and procedures section (Section 3.1.3) in greater detail.

All participants were able to participate using computers at the Assessment Cognitive Ergonomics Lab (Whittemore Hall room 519-A) at Virginia Polytechnic Institute and State University campus, at their homes, or any other convenient location. However, all of the participants chose to participate from computers at their homes, apartments, or dormitory rooms.

3.1.2 Equipment / Questionnaire

No equipment was necessary in this study except a computer connected to the Internet. There were no hardware or software specifications for the computer that could be used since the tasks did not require any multimedia or graphic applications. The

participants were also allowed to use any browser (e.g., Internet Explorer, Netscape, AOL).

The ONAT was in the form of a Web-based questionnaire (see Appendix E for Prototype B). Since Web-based questionnaires are devices that can be used for conducting remote usability evaluations, the study was a remote usability evaluation. Remote usability evaluations are defined as evaluations, where the evaluator(s) and the user(s) participate at evaluation session(s) at different locations and/or times (Hartson, Castillo, Kelso, Neale & Kamler, 1996). Even though the remote usability may have low internal validity, Smith-Jackson and Williges (2000) have suggested that these evaluations have high ecological validity, allowing the users to work in their natural environment. In other words, remote usability studies will not enable researchers to draw powerful conclusions or inferences about the cause and effects due to low internal validity (Sutherland, 1995a) but they will enable the results to be generalized to ordinary life rather than just certain experimental situations due to high ecological validity (Sutherland, 1995b).

The online tool consisted of three different types of questions:

1) **Needs Elicitation Questions:** Needs elicitation questions were directed to elicit the healthcare information needs of users. These questions consisted of general questions on computers, Internet usage and content preferences regarding healthcare websites. In Part I, a total of four 9-point Likert-Type scale questions (Questions 6, 7, 8, and 9) were used. For each question, a 9-point rating scale and its anchor explanation (as shown in Figure 2) were used to elicit responses.

(1 - Extremely Unimportant; 3 - Unimportant; 5 - Neutral; 7 - Important; 9 - Extremely Important)

1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○

Figure 2: 9-Point Likert-Type Scale

Needs elicitation questions also consisted of closed-ended questions with check-box/radio button response alternatives, open-ended questions, and a section for comments and suggestions (see Appendix F for sample questions).

2) Usability Questions: Usability Questions were directed to evaluate the usability of the ONAT. These questions allowed users to rate the online assessment tool they are completing on content, navigation and selection tools, information design and layout, and other necessary features. The usability questions were in the form of 9-point Likert-type rating scales, in order to quantify the user ratings of satisfaction, ease of use, comprehension and content importance. In Part II, a total of four 9-point Likert-Type scale questions (Questions 1, 2, 3, and 4) were used. For each question, a 9-point Likert-Type scale (see Figure 2 for the scale) and one of the two different anchor explanations were used to elicit responses. The first two of the Likert-type scale questions (Question 1 and 2) were used to have participants rate the difficulty level of certain tasks they completed in Part I. For each of these questions, the tasks were represented with response alternatives and the following rating scale anchors were used:

1 - Very Difficult; 3 - Difficult; 5 - Neutral; 7 - Somewhat Easy; 9 - Very Easy.

The last two of the Likert-type scales questions (Question 3 and 4) were used to have participants rate the general usability statements about the questionnaire they completed in Part I. For each of these questions, the usability statements were represented with response alternatives and the following rating scale anchors were used:

1 - Strongly Disagree; 3 - Disagree; 5 - Neutral; 7 - Agree; 9 - Strongly Agree.

In addition, a section enabled the users to communicate any comments or suggestions they had regarding usability of the questionnaire (see Appendix G for sample questions).

3) Guidelines Evaluation Questions: Guidelines evaluation questions were directed to evaluate the success rate of each usability guideline implementation and to determine the importance of implementing each individual guideline. In Part III, a total of fourteen 9-point Likert-Type scale questions (Questions 1 to 14) were used. For each of the guideline implemented in the tool, two separate 9-point Likert-type rating scales (see Figure 2 for the scales) were provided for users to rate the implementation importance of each guideline and the degree to which each guideline was successfully implemented (see Appendix H for sample questions). The first scale was identified as agreement scale and the second scale was identified as Importance scale. For both scales the following anchors were used:

1 - Strongly Disagree; 3 - Disagree; 5 - Neutral; 7 - Agree; 9 - Strongly Agree.

The agreement scale was used to have the participants assess their agreement with major characteristics of Prototype B, which were presented with an

individual guideline for each question. The importance scale was used to have the participants assess the importance of major characteristics of Prototype B, which were presented with the same guideline for each question. Each guideline was presented in the form of a statement where participant ratings determined the importance of each guideline and its' implementation success. In addition, a section enabled the users to communicate any comments or suggestions they had regarding the usability guidelines.

To prevent older users from confusion and frustration (see Section 2.1.2 for more details), the ONAT did not incorporate or use any embedded auditory cues but incorporated help aids. Also, the tasks in the online tool were not limited by time constraints. However, the each user was asked to record the duration it took to complete each section. Upon the completion of the questionnaires, the users were asked to report these times, which were used as one of the variables, to the experimenter.

3.1.3 Research Design and Procedures

The overall flow of events for Study 1 is illustrated in figure 3 on the next page.

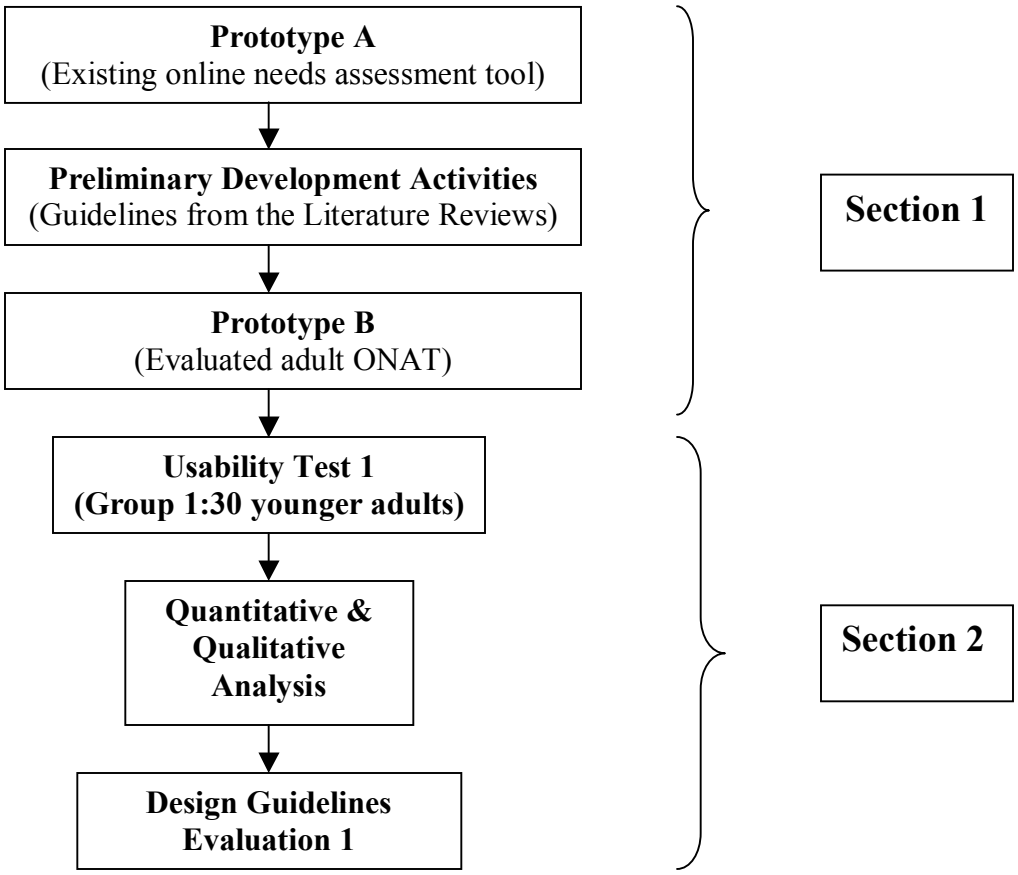


Figure 3: The overall flow of events for Study 1.

This particular study was an extension of ethnographic research, which is described as the study of participants in their natural environment without controlling any variables or making any manipulations (Sutherland, 1995c). In the previous research conducted by Smith-Jackson et al. (2001), to develop online assessment tools for healthcare information systems, content based, interactive focus groups and remote evaluations were conducted. The data gathered from these sessions supported the development of guidelines that were derived from the older adult users' needs and requirements (see Appendix C for the summary of results). These guidelines were implemented to construct an ONAT (Prototype A), which was also tested with older

users. However, during this research, Prototype A has not been revised or improved according to the need expressed by the results of the user testing.

Therefore, the first section of Study 1 was a heuristic evaluation by using guidelines extracted from existing research results (see Appendix I for the summary of guidelines) and the results of the previous study of Smith-Jackson et al. (2001). In this section, the experimenter (who was the expert in the study) conducted his evaluation guided by "guidelines, general principles and general rules of thumb", which is described as heuristic evaluation (Nielsen and Molich, 1990). At the end of section 1, the ONAT that was constructed and tested (Prototype A) in the previous research was improved with preliminary development activities. Subsequent to the preliminary development activities, the revision and evaluation of Prototype A was made according to the results of the heuristic evaluation. Based upon these evaluations and revisions, Prototype A was modified to form Prototype B (see Appendix K for the summary of heuristic evaluation steps). The summary of the design guidelines set used to develop Prototype B is as follows:

Color, Background, Foreground

1) Maximize contrast between characters and background.
2) Use easily discriminated color combinations.
3) Avoid hard coding on font, size and color.
4) Avoid distracting background elements.
5) Avoid animation.
6) Avoid exceptionally bright, fluorescence or vibrant colors.
7) Avoid blue/yellow or red/green combinations.
8) Maximize contrast between foreground and background.
9) Design in black and white; add color for emphasis.
10) Don't rely on color alone for information cue.
11) Maximize color contrast for people with color deficiencies.
12) Label each page with the site and page name.
13) Background color should be void of the blue/green range.
14) Graphics should be kept to a minimum or voided altogether.

Font

1) Use legible, commonly-used font throughout the site.
2) Use 12 to 14-point font sizes, and two points larger for titles.
3) Avoid mixed and decorative fonts.
4) Use bold type weight to emphasize.
5) Avoid all capital letters even in headings.
6) Use medium level type weight.
7) Use a sans serif font, such as Arial or Helvetica.

General Navigation Tools

1) Clearly label keys.
2) Use meaningful icons and label them.
3) Area cursor and sticky icon interaction techniques should be used at the same time.
4) Scrolling to view information should be minimized. Horizontal scrolling should be avoided.
5) To assist with navigation, users should be given pointers to get to special functions on the Web site. For instance, if a Web site allows users to automatically email information to a friend, that function should be highlighted with a call-out or instruction box with brief text on how to perform that function.
6) Buttons used to navigate, submit, or activate must be easy to view. Redundant coding using size and color enhance usability.
7) The online assessment tool instrument used a navigation button marked “Next” at the bottom of each page. “Next” and other navigation buttons should be used whenever possible.

Information Design and Terminology

1) Minimize irrelevant information.
2) Present information in consistent locations.
3) Adhere to principles of perceptual organizations.
4) Highlight important screen information.
5) Avoid complex command languages.
6) Use consistent operating procedures.
7) Balance the use of open space.
8) Leave right margin for various window/display sizes.
9) Avoid reduced kerning or condensed spacing.
10) Ensure that documents are clear and simple.
11) Use the same style throughout the site.
12) Leave enough space between paragraphs.
13) The visual field should be placed in narrow view.
14) Distorted spacing should be avoided; the block justification style should not be used.

15) Design your application consistent with software applications commonly used.
16) Plan screen sizes to fit in lower resolution display setting, at least 800*600 and perhaps even 640*480 for monitors up to a 17-inch screen size. Plan for users to browse with a single maximized (full-screen) window, and minimize the use of links and references to new browser windows.
17) Information layout should avoid clutter or “business”.
18) Information shall be presented in a column format to avoid the need to scroll horizontally.
19) In addition to avoiding clutter, links, menus, and function or navigation links embedded in a page shall be easily distinguishable from other text.

Links

1) Avoid link names like “click here”.
2) Avoid links requiring precise mouse control.
3) Make sure that links (a) are placed where they are easy to see, (b) are fairly large (the size settled on by the PD team here was 180 pixels * 22 pixels for a graphic button), and (c) have plenty of dead space around them to prevent accidental selection.

Help & Online/Offline-Support

1) Provide easy to use help.
2) Provide alternative mechanisms for on-line forms.
3) Simple navigation tools or tips should be provided to users when they enter the portal page.
4) The online tool must be tested in all browser environments to ensure usability.
5) Developers should be aware that many participants might complete the survey during non-business hours. Thus, some type of technical support during these hours may be needed.
6) Use open fields, check boxes, and radio buttons as needed.
7) Minimize the need to enter information in fields. Reduce typing demands.

Question Content and Structure

1) Questions related to information density should refer to “clutter” to assess user’s opinions of information density.
2) Questions referring to “ease-of-use” or “simplicity” are easily understood by users and should be included as much as possible to assess design needs and preferences.
3) Questions and response alternatives should be developed using information from focus groups.
4) “Other” or open fields should be used for questions with response alternatives that are not comprehensive.

The needs, requirements, and expectations of younger adults were determined by the findings of test group 1 (in Section 2 of Study 1). The process of selecting and assigning the participants to the test groups was based on two different criteria. The first criterion required 10 participants or more for test group 1 and the second criterion required equal numbers of participants for each category (categories being novice, intermediate and expert users) in test group 1. According to these criteria, the usability test group 1 consisted of a total of 30 participants (10 for each category). The reason for selecting a sample size of 10 participants or more was due to the finding that during interface design and testing, 90 percent of the problems related to usability were found to be detected by using a sample size of 10 participants (Virzi, 1992). In regard to this finding, the sample sizes were sufficient to capture most usability problems.

During the assignment of participants, 40 volunteers were recruited. Every volunteer was given the participant categorization tool. However, they were only assigned as a participant if there was a need for participants in their experience level category. The recruitment was carried on until all 30 of the Test Group 1 participants were assigned.

After the assignment, the participants were provided with one of the two versions of informed consent forms according to the age group they were in (see Appendix L and Appendix M for the two different versions). These forms were sent either via email as attachments or provided as hardcopies in person. The participants were asked to read and type their names in the space provided under participant permission section of the informed consent form. By typing their names, the participants communicated their consent to participate in the study. After receiving their consent, the participants were

be given a user ID number, a password and the Uniform Resource Locator (URL) of the Online Demographic Information Questionnaire so that they were be able to log on and complete the online questionnaire (see Appendix N for the questionnaire). The Demographic Information Questionnaire was used to gather demographic information of the participants (see Appendix O for the summary of participant demographic information) and had the same format as Prototype B (the ONAT that was tested with Usability Test 1). Using the same format allowed the participants to use this initial questionnaire, as a practice session to become familiar with the format of the actual ONAT and the identification of potential technical problems such as computer or browser incompatibility. After eliminating the identified technical problems, the URL of Prototype B was distributed to the participants for them to log on and complete the assessment tool. The user ID number and password assignment also helped control the actual users who completed the questionnaire and Prototype B. The user ID number and passwords were active for 2 weeks to provide enough time for the participants to complete the questionnaire and Prototype B. The participants were also provided with the experimenter's telephone number and email address in case problems or technical difficulties occur. At the end of 2 weeks, the user ID numbers and the passwords expired. Upon the completion of Prototype B, pressing the submit button allowed the answers to be collected and emailed to the researcher. After the participant submitted the responses, younger adult users were compensated with their preference of either $7.50 or 1 research credit.

3.1.4 Statistical Analysis

The data analyses were conducted with the quantitative and qualitative raw data gathered from usability test group 1. The qualitative data were gathered from the answers to the open-ended questions of the Needs Elicitation Questions and from the responses to the section for comments and suggestions of all of the three question types (see section 3.1.2 for details). The quantitative data were gathered from the answers to the closed-ended questions with check box and radio-button (9-point Likert-Type scales) alternatives of the Needs Elicitation Questions, the usability ratings of the Usability Questions and the importance and implementation success ratings of the Guideline Evaluation Questions (see Section 3.1.2 for details).

The quantitative data were analyzed by comparing ratings of users across each independent variable to determine significance. The answers to the closed-ended questions were treated as nominal data and were analyzed by using a Chi Square Test. The results of radio-button (9-point Likert-Type scales) alternatives of the Needs Elicitation Questions, usability and guideline evaluation ratings (second and third type of survey questions) were treated as ordinal data and were analyzed by using an ANOVA test. The decision for treating the usability ratings as ordinal data was made based upon the small sample size (a sample size of 10 participants) of test group 1 and non-parametric tests being statistically more powerful for samples sizes less than 30 (Reaves, 1992). However, instead of Kruskal-Wallis Median tests, which would be appropriate for the treatment conditions and the data type, ANOVA tests were conducted due to the limitations of the parametric tests with SAS 6.12 software. This decision was made due to the non-parametric one-way analysis procedure treating all

the medians equal and its' inability to conduct multiple comparisons. Also, as non-parametric one-way analysis procedure is the equivalent to the one-way ANOVA procedure, which uses Chi-square estimation values and asymptotic p values, this substitution is commonly practiced. In addition, the results of completion times were analyzed by using a separate ANOVA test. For all of the separate statistical analyses performed, the dependent variables were the participant responses to the ONAT Prototype B questions and the completion time. The independent variable was the experience categories (3 levels). Subsequent to the ANOVA tests performed, tests of reliability were performed to determine the internal consistency of the response alternatives. For each of the response alternative, the alpha reliability was measured using Cronbach's alpha reliability measure. The Cronbach coefficient alpha was used to calculate the average correlation among the response alternatives that were collected at the same time. The average correlation was then used to estimate the reliability or the internal consistency of the scale (Garson, 2004).

The qualitative data were analyzed by using content analyses. Content analysis, which is the analysis that engages in frequency counting of chosen words or phrases in a context (Wilson and Corlett, 1995), was performed by the experimenter. During the content analysis, for each subsequent question, phrases were gathered to create the initial response categories. After gathering the initial response categories, the categories were reorganized, refined and merged to develop the final response categories. The first level of content analysis was finalized by performing frequency counts for each of the final response categories. A more in-depth second level of content analysis that would track the benchmarks such as prominence (the words

specifically positioned or highlighted to direct attention to a phrase), off-spin (amplified or misinterpreted words or phrases), wordplay (combination or words to invoke a strategic verbal exchange), verbatims (corresponding with word for word), positioning (the arrangement of words or phrases in a sentence or in a text to express a thought or an opinion), and notational visibility (the usage of fonts to direct attention) (Solomon, 1993) was not conducted due to the insufficient amount of qualitative data.

According to results of the statistical analyses, no new set of guidelines were extracted and added to the current set of guidelines. The guideline extraction was conducted in three different sections. Four different methodologies and sets of criteria were used for extracting the guidelines for each of the different section.

In the first section, qualitative results were used to extract guidelines. The qualitative results consisted of the responses to the open-ended questions of the Needs Elicitation Questions and of the responses to the section for comments and suggestions of the three question types (see Section 3.1.2 for details. After gathering all the qualitative responses, each of the response was classified in one of the 5 categories shown in Table 1.

Table 1: The Table of Classification Categories.

Categories	Importance Values
1) Wishes	1
2) Suggestions	2
3) Concerns	3
4) Wants	4
5) Needs/Requirements	5

The classification was performed considering the respondent's language, word selection, indications of stress and punctuation. The importance values for each of the five categories provided in Table 1, were generated by applying the quantification approach of Force Field Analysis. Force Field Analysis, which is a decision-making technique proposed by Weisbord (1987), is an analysis method used in order to quantify the pros and cons of a change or an innovation. This analysis method uses a simple quantification approach where quantitative values are assigned to qualitative data by the researcher in order to enable decision-making.

After classifying each response for the particular question in categories, the number of responses within each category was counted and the overallsScore was calculated (according to the formula below).

$$\textbf{The Overall Score} = \frac{\sum_{1}^{n} (The\ \#\ of\ responses\ for\ category\ n) * (n)}{The\ \#\ of\ total\ responses\ for\ the\ question}$$

n = The 5 different categories from (1 to 5)

Subsequent to calculating the overall score for the specific question, the following criteria were used in order to determine whether or not the response should be developed as a design guideline:

a) If the overall score exceeds 0.4, the specific response will be developed into a guideline.
b) Any guideline that cannot be implemented (due to technical restrictions) will be ignored.

This method was used for the responses gathered from each of the questions individually.

In the second section, the first segment of the quantitative results was used to extract guidelines. The first segment of quantitative results consisted of the responses to the closed-ended questions (with radio-button and check-box alternatives) of the Needs Elicitation Questions. Subsequent to gathering the first segment of quantitative responses, the following criteria were used in order to determine whether or not the response should be developed as a design guideline:

a) If the majority of the respondents agree (60% or more of the participants) that a certain response alternative is unimportant or insignificant, the response alternative will be excluded from the design and a supporting guideline will be developed.

b) If the majority of the respondents do not agree (less than 60% of the participants) that a certain response alternative is unimportant or insignificant, no changes will be made and no guideline will be developed.

c) Any guideline that cannot be developed or implemented (due to technical restrictions) will be ignored.

This method was used for the responses gathered from each of the questions individually.

In the third section, the second segment of the quantitative results was used to extract guidelines. The second segment of quantitative results consisted of the responses to the ratings of the usability questions. Subsequent to gathering the first segment of quantitative responses, the following criteria were used in order to determine whether or not the response should be developed as a design guideline:

a) If the majority of the respondents rate (the mean of the ratings is 6 or more, since the rating “6” presents the anchor “agree”) their satisfaction with the task or the item, no guidelines will be developed.
b) If the majority of the respondents do not rate (the mean of the ratings is less than 4, since the rating “3” presents the anchor “disagree” and the rating “5” represents “neutral”) their satisfaction with the task or the item, a guideline will be developed in order to avoid the dissatisfaction.
c) Any guideline that cannot be developed or implemented (due to technical restrictions) will be ignored.

This method was used for the responses gathered from each of the questions individually.

In the last section, the third segment of the quantitative results was used to extract guidelines. The third segment of quantitative results consisted of the responses to the implementation success ratings and guideline importance questions of the Guideline Evaluation Questions. Subsequent to gathering the second segment of quantitative responses, the following criteria were used in order to determine whether or not the response should be developed as a design guideline:

a) If the majority of the respondents rate (the mean of the ratings is less than 4, since the rating “3” presents the anchor “disagree” and the rating “5” represents “neutral”) that a guideline is not important on importance scale, the guideline’s implementation will be excluded from the design.
b) If the majority of the respondents rate (the mean of the ratings is 6 or more, since the rating “6” presents the anchor “agree”) that a guideline is important on

importance scale, the guideline's evaluation will be based on the agreement scale results.

c) If the majority of the respondents rate (the mean of the ratings is 6 or more, since the rating "6" presents the anchor "agree") that a guideline is not implemented successfully on agreement scale, alternative implementation methods will be executed within technical restrictions.

This method was used for the responses gathered from each of the questions individually.

According to the three different methodologies and set of criteria mentioned, no guidelines were extracted and added to the current set of guidelines. Therefore, the Design Guidelines Evaluation 1 only consisted of the guideline set used to develop Prototype B.

3.1.5 Results

This section represents the results of the quantitative and qualitative data analysis gathered from Usability Test Group 1. In the following segment, the results of the qualitative and quantitative data are presented separately for each questionnaire part.

Part I

Part I consisted of Needs Elicitation Questions in the form of: a) Closed-ended questions with check-box alternatives, b) Closed-ended questions with radio-button alternatives, and c) Open-ended questions (see Section 3.1.2 for more details). The results for each of these different question forms are presented in Appendix J.

Part II

Part II consisted of Usability Questions in the form of: a) Close-ended questions with radio button alternatives and b) Comments and suggestions section (see Section 3.1.2 for more details). The results for each of these question forms are as follows:

<u>a) Close-Ended Questions with radio-button alternatives (9-point Likert-Type scale questions):</u>

The rating distribution chart for the task difficulty levels is shown in Figure 4. As indicated, 94% of the participants provided an overall rating of "7" or above and agreed that the tasks they completed were easy.

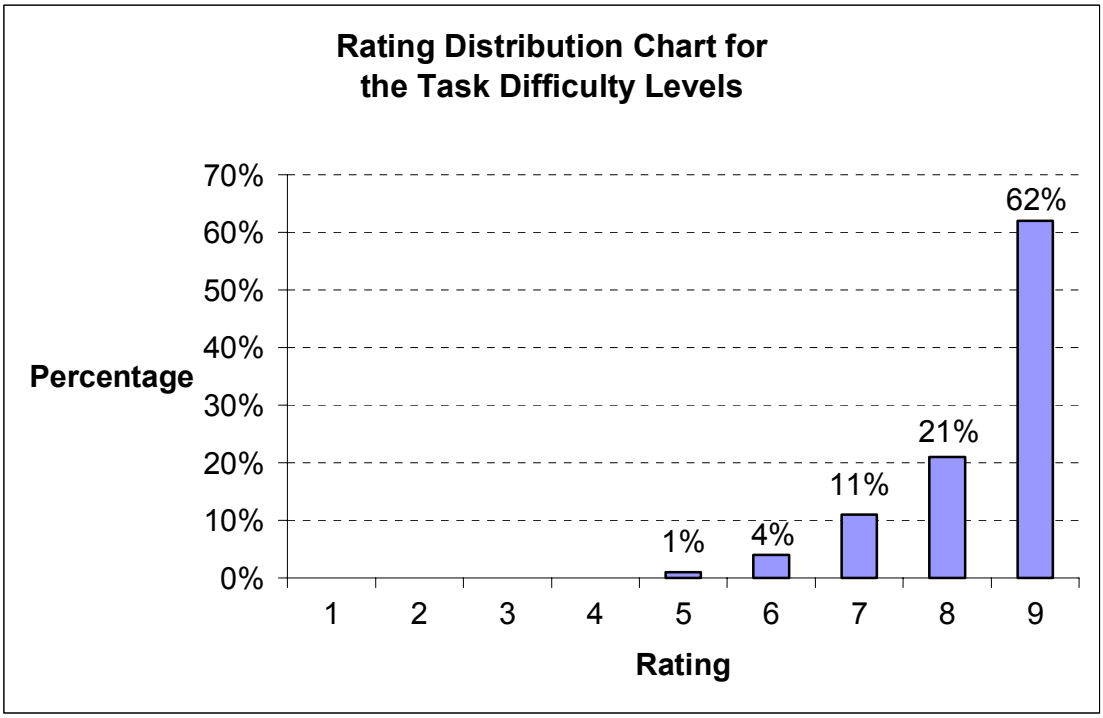


Figure 4: Rating Distribution Chart for the Task Difficult Levels

The mean ratings and overall rating distribution for each task are provided in Table 10.

Table 10: Task Difficulty

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Ratings	SD
Selecting response circles	0	0	100	8.90	0.03
Going to the next page	0	0	100	8.60	0.06
Understanding the concept of an "Online Questionnaire"	3.3	0	96.7	8.56	0.06
Selecting responses squares	0	3.3	96.7	8.53	0.05
Operating my input devices	0	3.3	96.7	8.50	0.05
Filling in the blanks	0	3.3	96.7	8.30	0.26
Reading the questions	3.3	10	86.7	8.20	0.33
Opening or getting into the survey	3	7	90	8.03	0.51
Understanding what is being asked	0	13.3	86.7	7.80	0.98

The rating distribution chart for the general usability statements is shown in Figure 5. As indicated, 88% of the participants provided an overall rating of “7” or above and agreed that the Online Tool was easy to use, sufficient, adequate, and concise.

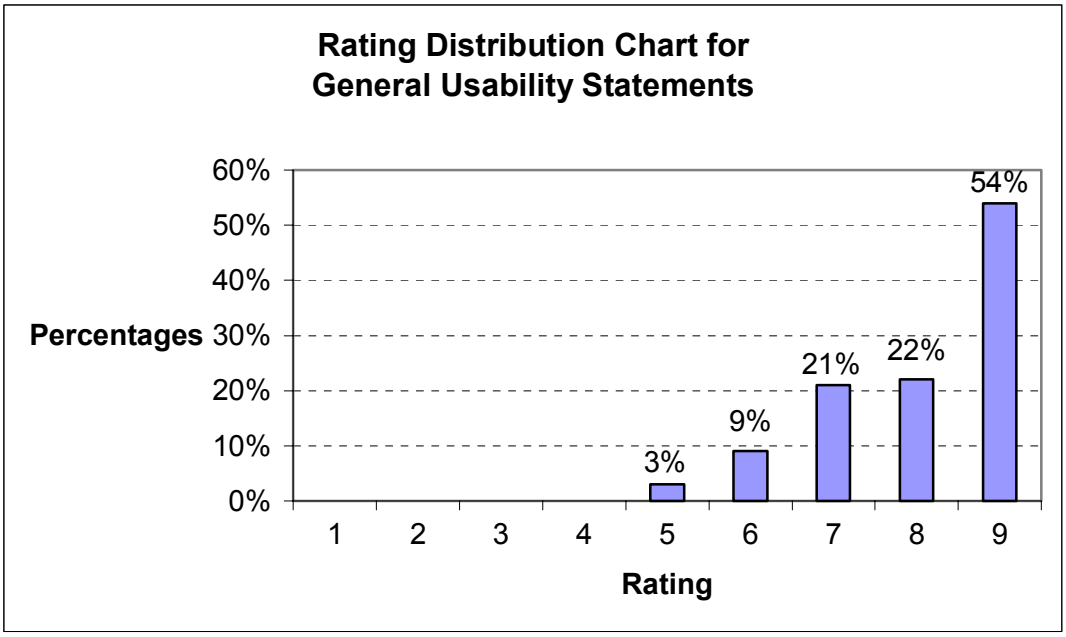


Figure 5: Rating Distribution Chart for the General Usability Statements

The mean ratings and overall rating distribution for each statement are provided in Table 11.

Table 11: Usability Statements

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Ratings	SD
Overall, the amount of effort it took to complete this questionnaire was acceptable for me	**0**	**3.3**	**96.7**	**8.53**	**0.05**
This Online Questionnaire is easy to use	**3.3**	**3.3**	**93.4**	**8.46**	**0.09**
I am comfortable with online evaluations	**0**	**3.3**	**96.7**	**8.36**	**0.1**
I was comfortable with the amount of questions in the Online Questionnaire	**3.3**	**6.6**	**90**	**8.33**	**0.1**
I was comfortable with the content of questions in the Online Questionnaire	**0**	**0**	**100**	**8.26**	**0.18**
Overall, the amount of time it took to complete this questionnaire was acceptable for me	**3.3**	**10**	**86.7**	**7.96**	**0.26**
I was comfortable with the length of the Online Questionnaire	**3.3**	**10**	**86.7**	**7.96**	**0.33**
The questions in the online questionnaire were adequate for me to communicate all my needs and preferences	**6.6**	**6.6**	**86.7**	**7.90**	**0.51**
It was easy to view the Online Questionnaire as a link while completing this form	**3.3**	**3.3**	**93.4**	**7.90**	**0.63**
Using the Online questionnaire helped me communicate what I want and don't want in a Web site	**0**	**13.3**	**86.7**	**7.80**	**0.98**
I would prefer to meet face-to-face with a researcher to evaluate The Online Questionnaire	**6.6**	**36.6**	**53.3**	**6.56**	**1.23**

Subsequent to calculating the mean ratings, the guideline extraction methodology was applied to the results of these four questions. However, neither one of the response alternatives met the guideline extraction criteria (see Section 3.1.4 for more details on the criteria). Therefore, no guideline was developed.

In addition to the guideline extraction, 2*3 ANOVA was performed for each of the response alternatives and completion times. Each ANOVA test was performed for experience level categories and gender (see Appendix R for ANOVA results). The results of these analyses indicated only one significantly different result.

For the usability statement “Selecting response squares” (response alternative 4 of question 1), the following significant effect was found:

- There was a significant difference between the mean rating value for female and male participant groups ($F\ (1,24) = 8.47,\ p < .05$). Tukey’s Studentized Range (HSD) test indicated that the mean rating score for female participants ($M = 8.93,\ SD = 0.07$) was significantly higher than the mean rating score for male participants ($M = 8.13,\ SD = 0.98$). Female participants found selecting response squares significantly easier to use than male participants.

The Cronbach coefficient alpha value was *.77* and the range of item-to-total correlation was *.04 – .78*. The correlation coefficients among the 20 response alternatives were all positive and statistically significant at the *.05* level. The correlation coefficients ranged from a low *.04* to a high *.78*, indicating the degree of the relationships between the response alternatives. The lowest reliable variance of the total scores was at least 77%. Each response alternative had adequate internal consistency and there was no need to further improve the reliability of these questions and their response alternatives.

b) Comments and suggestions sections:

In this Section, participants were asked to provide feedback regarding the questionnaire, participation process, needs, requirements and preferences. For the results gathered from this section, content analysis was performed. Comments and suggestions were identified as final response categories (as shown in Table 13) and guideline extraction methodology was applied.

Table 13: Qualitative Responses

Responses		Frequency
Comments	Questionnaire was short and straightforward	3
	Questionnaire asked all the right questions	3
	Questionnaire was complete	4
Suggestions	The next button should not be so far down on the page	1
	It may not be necessary to have response circles ranging from 1-9. Maybe only 1-4 is enough?	1

The responses for comments category consisted of phrases participants used to communicate their satisfaction with the tool. The responses for the suggestions category consisted of two design alternatives. However, neither one of these alternatives exceeded the overall score 0.4 (see Section 3.1.4 for more details on calculating the score). Therefore, no new design guideline was extracted for these two response categories.

Part III

Part III consisted of Guideline Evaluation Questions in the form of: a) Close-ended questions with radio button alternatives and b) Comments and suggestions section (see Section 3.1.2 for more details). The results for each of these question forms are as follows:

<u>a) Close-Ended Questions with radio-button alternatives (9-point Likert-Type scale questions):</u>

The Rating Distribution Chart for Agreement Scale is shown in Figure 6. As indicated, 84% of the participants provided an overall rating of “7” or above and agreed that each design guideline of Prototype B was well implemented.

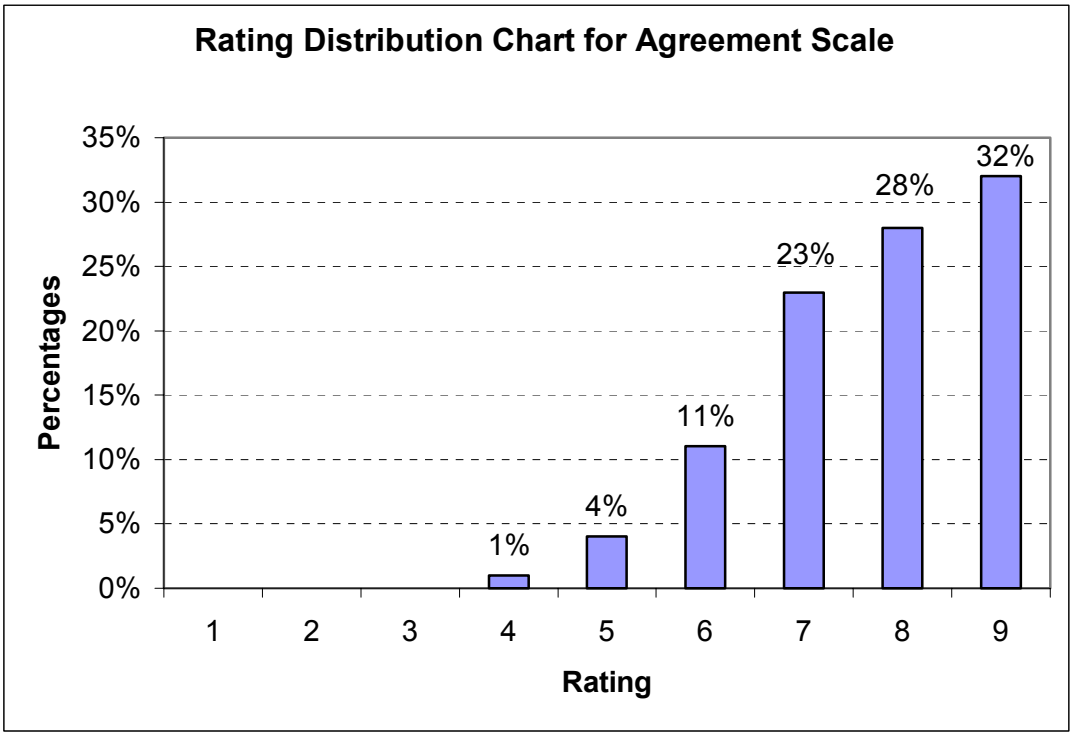


Figure 6: Rating Distribution Chart for Agreement Scale

Agreement scale mean ratings and overall rating distribution for each guideline are provided in Table 14

Table 14: Agreement Scale Mean Ratings and Overall Rating Distribution

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Ratings	SD
Open fields, check boxes, and radio buttons are sufficient. Typing demands are reduced	**0**	**3.3**	**96.7**	**8.10**	**0.22**
Information layout is not cluttered. Language is clear, simple and not technical	**3.3**	**6.6**	**90**	**8.06**	**0.24**
Visual field is placed in narrow view, the use of open space is balanced and sufficient right margin is left for various window/display sizes	**3.3**	**10**	**86.7**	**7.96**	**0.26**
Provide easy to use help during office and non-business hours	**13.3**	**16.6**	**70**	**7.96**	**0.26**
The Online Questionnaire works properly on your computer and your browser	**3.3**	**13.3**	**83.4**	**7.73**	**0.42**
Graphics, distracting background elements and animation are kept to a minimum or voided altogether	**3.3**	**13.3**	**83.4**	**7.73**	**0.42**

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Ratings	SD
Links, buttons, keys and icons are of proper size and color, clearly labeled, and meaningful	**3.3**	**3.3**	**93.4**	**7.70**	**0.42**
Links, buttons, keys and icons are sufficient, easily distinguished from other text and easy to access	**3.3**	**10**	**87.6**	**7.66**	**0.51**
Page layout (placement or spacing of headings, types, paragraphs, labels or other page elements) is properly designed	**6.6**	**10**	**83.4**	**7.66**	**0.58**
The information presented is relevant	**3.3**	**16.6**	**80**	**7.63**	**0.58**
Font, typeface (for example, bold, italic, underline, capital letters etc.), and size are properly applied for the text	**6.6**	**9.9**	**83.4**	**7.56**	**0.65**
Scrolling to view information is minimized. Horizontal is avoided	**6.6**	**13.3**	**76.7**	**7.20**	**0.71**
Information presentation and operating procedures are logical and consistent with other commonly used applications and throughout the Online Questionnaire	**0**	**3.3**	**97.6**	**7.10**	**0.33**
Color is properly used for the characters, foreground, background, and to highlight important information	**10**	**30**	**60**	**7.10**	**0.99**

The Rating Distribution Chart for Importance Scale is shown in Figure 7. Ninety–Three percent (93%) of the participants provided an overall rating of “7” or above and agreed that each design guideline of Prototype B was important.

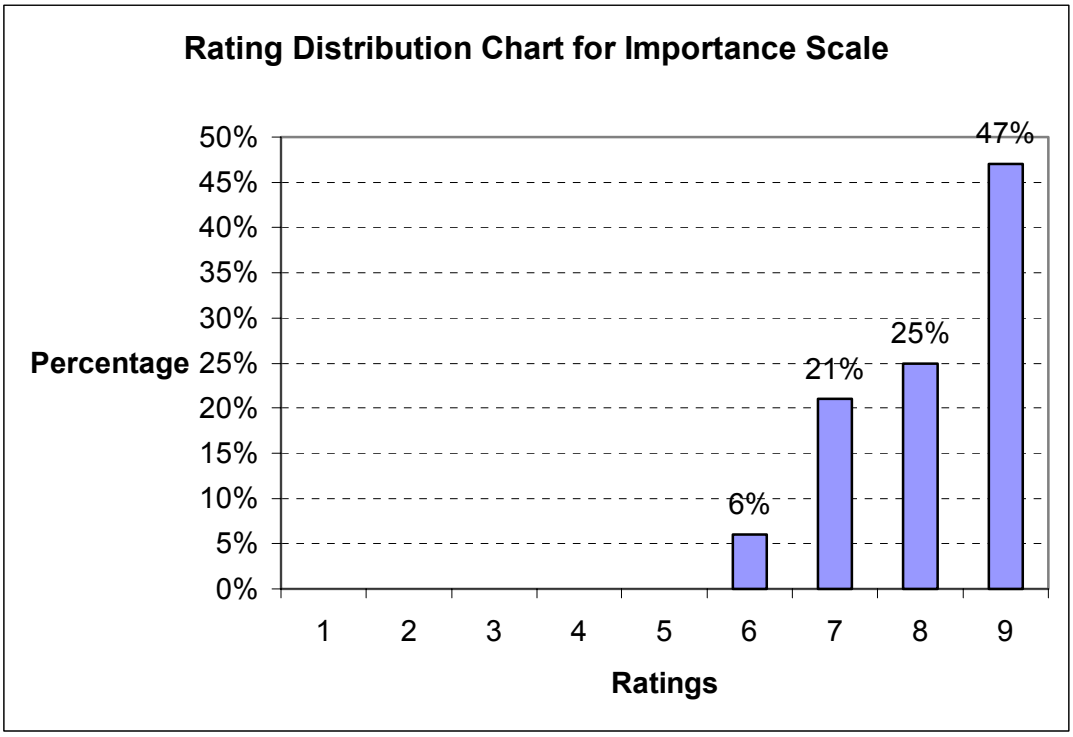


Figure 7: Rating Distribution Chart for Importance Scale

Importance scale mean ratings and overall rating distribution for each guideline are provided in Table 15.

Table 15: Importance Scale Mean Ratings and Overall Rating Distribution

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Rating	SD
Provide easy to use help during office and non-business hours	**0**	**0**	**100**	**8.60**	**0.03**
Scrolling to view information is minimized. Horizontal is avoided	**0**	**0**	**100**	**8.36**	**0.06**
Information presentation and operating procedures are logical and consistent with other commonly used applications and throughout the Online Questionnaire	**0**	**0**	**100**	**8.30**	**0.06**
Information layout is not cluttered. Language is clear, simple and not technical	**0**	**10**	**90**	**8.30**	**0.12**
Color is properly used for the characters, foreground, background, and to highlight important information	**0**	**6.6**	**93.4**	**8.20**	**0.12**
Open fields, check boxes, and radio buttons are sufficient. Typing demands are reduced	**0**	**13.3**	**86.7**	**8.20**	**0.21**

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Rating	SD
The information presented is relevant	**0**	**10**	**90**	**8.16**	**0.11**
Visual field is placed in narrow view, the use of open space is balanced and sufficient right margin is left for various window/display sizes	**0**	**6.6**	**93.4**	**8.13**	**0.08**
Links, buttons, keys and icons are sufficient, easily distinguished from other text and easy to access	**0**	**3.3**	**96.4**	**8.13**	**0.06**
Font, typeface (for example, bold, italic, underline, capital letters etc.), and size are properly applied for the text	**0**	**3.3**	**96.7**	**8.13**	**0.05**
Page layout (placement or spacing of headings, types, paragraphs, labels or other page elements) is properly designed	**0**	**3.3**	**96.4**	**8.10**	**0.05**
Links, buttons, keys and icons are of proper size and color, clearly labeled, and meaningful	**0**	**6.6**	**93.4**	**7.86**	**0.26**
The Online Questionnaire works properly on your computer and your browser	**0**	**0**	**100**	**7.86**	**0.23**
Graphics, distracting background elements and animation are kept to a minimum or voided altogether	**3.3**	**13.3**	**83.4**	**7.46**	**0.51**

Subsequent to calculating the mean ratings, guideline extraction methodology was applied to the results of these fourteen guidelines. However, neither one of the response alternatives met the guideline extraction criteria (see Section 3.1.4 for more details on the criteria). Therefore, no guideline was developed.

In addition to the guideline extraction, 2*3 ANOVA was performed for each of the response alternatives and completion times. Each ANOVA test was performed for experience level categories and gender (see Appendix R for ANOVA results). The results of these analyses indicated that six guidelines showed significantly different results. For the agreement scale, the following significant effects were found:

- For Guideline 2 (Graphics, distracting background elements, and animation are kept to a minimum or voided altogether), there was a significant difference

between the mean rating value for female and male participant groups (*F(1,24) =12.23, p <.05*). Tukey's Studentized Range (HSD) test indicated that the mean rating score for female participants (*M =8.33, SD =0.98*) was significantly higher than the mean rating score for male participants (*M =7.13, SD =1.06*). Female participants found the Guideline 2 significantly better implemented than male participants.

- For Guideline 9 (Information presentation and operating procedures are logical and consistent with other commonly used applications and throughout the Online Questionnaire), there was a significant difference between the mean rating value for female and male participant groups (*F (1,24) =14.52, p <.05*). Tukey's Studentized Range (HSD) test indicated that the mean rating score for female participants (*M =8.60, SD =0.63*) was significantly higher than the mean rating score for male participants (*M =7.60, SD =0.83*). Female participants found the Guideline 9 significantly better implemented than male participants.

For Importance Scale, the following significant effects were found:

- For Guideline 3 (Font, typeface and size are properly applied for the text), there was a significant difference between the mean rating value for female and male participant groups (*F (1,24) =15.52, p <.05*). Tukey's Studentized Range (HSD) test indicated that the mean rating score for female participants (*M =8.66, SD =0.73*) was significantly higher than the mean rating score for male participants (*M =7.60, SD =0.82*). Female participants found the Guideline 3 significantly more important than male participants.

- For Guideline 4 (Links, buttons, keys and icons are of proper size and color, clearly labeled, and meaningful), there was a significant difference between the mean rating value for female and male participant groups ($F\ (1,24) = 19.64,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean rating score for female participants ($M = 8.46,\ SD = 0.74$) was significantly higher than the mean rating score for male participants ($M = 7.26,\ SD = 0.70$). Female participants found the Guideline 4 significantly more important than male participants.
- For Guideline 5 (Scrolling to view information is minimized. Horizontal is avoided), there was a significant difference between the mean rating value for female and male participant groups ($F\ (1,24) = 16.9,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean rating score for female participants ($M = 8.80,\ SD = 0.56$) was significantly higher than the mean rating score for male participants ($M = 7.93,\ SD = 0.59$). Female participants found the Guideline 5 significantly more important than male participants.
- For Guideline 8 (Links, buttons, keys and icons are sufficient, easily distinguished from other text and easy to access), there was a significant difference between the mean rating value for female and male participant groups ($F\ (1,24) = 7.78,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean rating score for female participants ($M = 8.53,\ SD = 0.74$) was significantly higher than the mean rating score for male participants ($M = 7.73,\ SD = 0.88$).Female participants found the Guideline 8 significantly more important than male participants.

- For Guideline 9 (Information presentation and operating procedures are logical and consistent with other commonly used applications and throughout the Online Questionnaire), there was a significant difference between the mean rating value for female and male participant groups ($F\ (1,24) = 9.29,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean rating score for female participants ($M = 8.73,\ SD = 0.59$) was significantly higher than the mean rating score for male participants ($M = 7.93,\ SD = 0.80$). Female participants found the Guideline 9 significantly more important than male participants.
- For Guideline 13 (The Online Questionnaire works properly on your computer and your browser), there was a significant difference between the mean rating value for female and male participant groups ($F\ (1,24) = 11.52,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean rating score for female participants ($M = 9.00,\ SD = 0$) was significantly higher than the mean rating score for male participants ($M = 8.20,\ SD = 0.86$). Female participants found the Guideline 13 significantly more important than male participants.

The Cronbach coefficient alpha value was *.88* and the range of item-to-total correlation was *.19* – *.64*. The correlation coefficients among the 20 response alternatives were all positive and statistically significant at the .05 level. The correlation coefficients ranged from a low *.19* to a high *.64*, indicating the degree of the relationships between the response alternatives. The lowest reliable variance of the total scores was at least 88%. Each response alternative had adequate internal consistency and there was no need to further improve the reliability of these questions and their response alternatives.

b) Comments and suggestions sections:

In this section, participants were asked to provide any other guideline that they thought was important and should be included in the design of the Online Questionnaire. For the results gathered from this section, content analysis was performed. Comments were identified as final response category (as shown in Table 17) and guideline extraction methodology was applied.

Two comments were identified as final response categories. A total of six participants commented that the guidelines were complete and adequate, whereas a total of five participants commented that they would not make any changes to any part of the questionnaire. As the responses for comments category consisted of phrases participants used to communicate their satisfaction with the guidelines and the online tool, no new design guideline was extracted for this response category.

In addition, since no new guidelines were extracted from the results of Study 1, the Design Guidelines Evaluation 1 consisted of the design guideline set (see Section 3.1.3 for the guideline set).

3.2 Study 2

As mentioned in the prior discussions, this study was designed to meet the secondary research objective, which was stated as: Whether and to what degree can one ONAT be developed to be usable for both older and younger users?

3.2.1 Participants

A total of 24 participants were selected from the local community. Half of these participants were selected from volunteers who were 50 years old or older and the rest of the participants were selected from volunteers who were between 18 and 35 years of age. The requirements regarding participant eligibility discussed in Study 1 also applied to Study 2. To form test groups 2 & 3 (as shown in Figure 3), all of the volunteer participants were given the participant categorization tool (see Appendix D). With the help of the categorization tool, the potential participants were be categorized into novice, intermediate, and expert users. The additional selection criteria for the participants will be discussed in the research design and procedures section (Section 3.2.3) in greater detail.

The research environments were also the same as in Study 1 (see Section 3.1.2 for details). In addition, participants were also able to participate from computers at the senior center at the Blacksburg Parks and Recreation Center. For test group 2, nine older participants completed the surveys at their house or apartments and the remaining three older participants completed the surveys at their work. For test group 3, ten younger participants completed the surveys at their house, apartments or dormitory rooms and the remaining two younger participants completed the surveys at their work.

3.2.2 Equipment / Questionnaire

The equipment requirements and questionnaire format were also identical to Study 1 (see Section 3.1.2 for details).

3.2.3 Research Design and Procedures

The overall flow of events for Study 2 is illustrated in Figure 8:

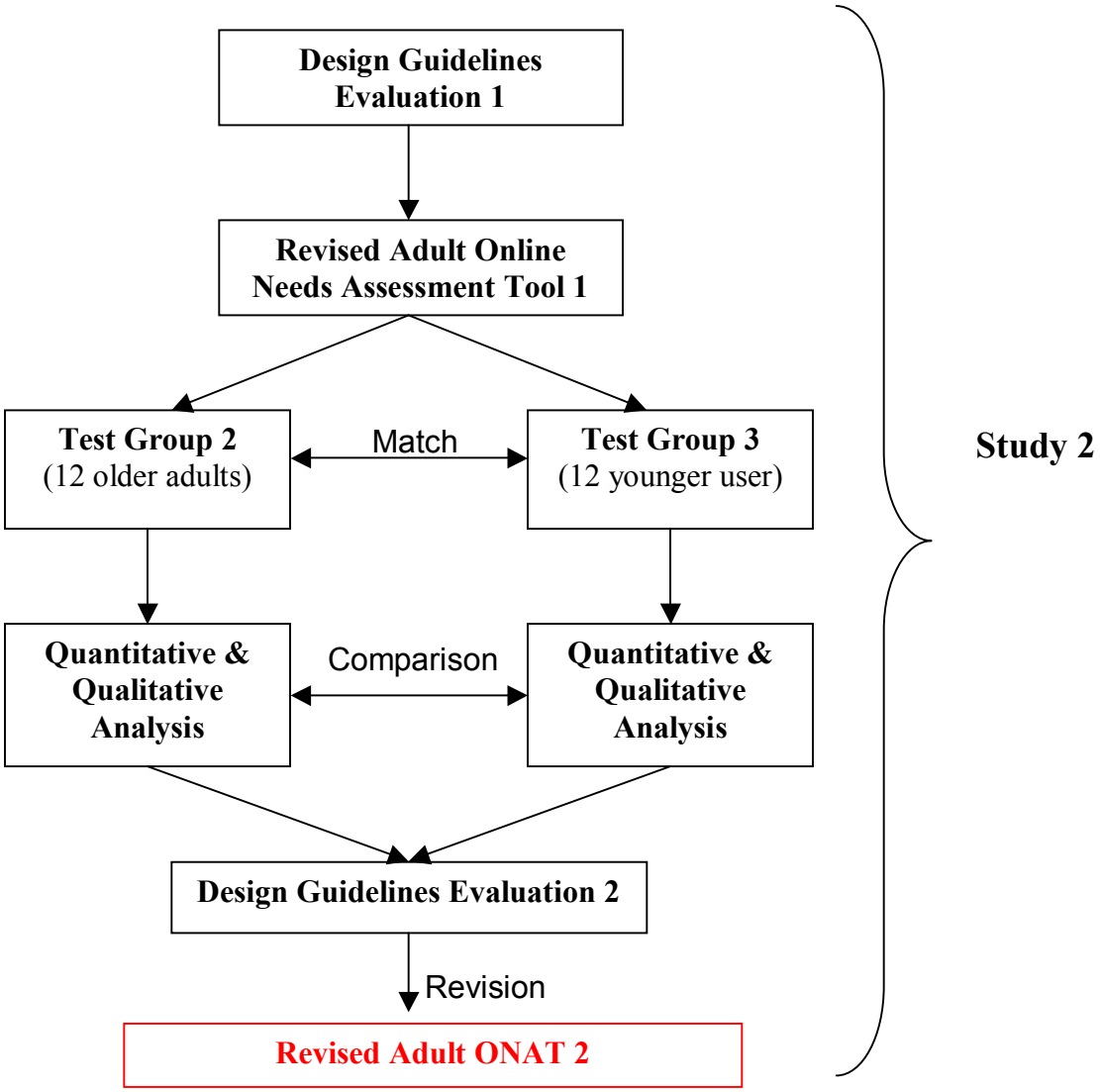


Figure 8: The overall flow of events for Study 2

Study 2 was an extension of study 1. In this study, the first design iteration was conducted in order to structure Revised Online Needs Assessment Tool 1. The design iteration consisted of implementing the Design Guidelines Evaluation 1 and revising the design of Prototype B to develop Revised Online Needs Assessment Tool 1. Even though Design Guideline Evaluation 1 consisted of the design guideline set, several changes and revisions were made to Prototype B (see Appendix Q for the summary of changes after Study 1)

To evaluate the usability of the Revised Online Needs Assessment Tool 1 developed for both older and younger users, test groups 2 & 3 were used. The process of selecting and assigning the participants to the test groups were based on four different criteria. The first criterion required 10 participants or more for each of the test groups, the second criterion required an equal numbers of participants for each category (categories being novice, intermediate and expert users), the third criterion required the assignment of older users to the test group 2 and younger users to test group 3 and the final criterion required the recruitment of as many participants as possible considering the limited number of older user volunteers in the local community. According to these criteria, test group 2 consisted of a total of 12 older user participants (4 for each category) and test group 3 consisted of a total of 12 younger users (4 of each category). As in Study 1, a sample size of 12 participants was selected on the basis of Virzi's (1992) finding that during interface design and testing, 90 percent of the problems related to usability were found to be detected by using a sample size of 10 participants. In regard to this finding, the sample sizes were sufficient to capture most usability problems.

During the assignment of participants, 31 volunteers were recruited. Every volunteer was given the participant categorization tool. However, they were only assigned as a participant if there was a need for participants in their age and experience level category. The recruitment was carried on until all 24 of the Test Group 2 and 3 participants were assigned.

After assigning the participants to the test groups, the same informed consent, online demographic information questionnaire administration, expiration, participant security protocols were used as in Study 1. Upon completing the questionnaires, younger users were compensated with their choice of $7.50 or 1 research credit and older users were compensated with $7.50.

3.2.4 Statistical Analysis

Data analyses were conducted with the quantitative and qualitative raw data gathered from the two test groups. The qualitative and quantitative data were gathered in the same way as used in Study 1 (see Section 3.1.4 for details). Also, the quantitative data were analyzed by comparing ratings to determine significance and the qualitative data were analyzed by content analysis.

The responses to the closed-ended questions were treated as nominal data and were analyzed by using a Chi Square Test. The results of radio-button (9-point Likert-Type scales) alternatives of the Needs Elicitation Questions, usability and guideline evaluation ratings (second and third type survey questions) were treated as ordinal data and were analyzed by using an ANOVA (see Section 3.1.4 for details). For both of the separate statistical analyses performed, the dependent variables were the participant responses to the final ONAT questions and completion time. The independent variables

were the test groups (2 levels), and the experience categories (3 levels)). In addition, tests of reliability were performed to determine the internal consistency of the response alternatives. For each of the response alternative, the alpha reliability was measured using Cronbach's alpha reliability measure. The answers to the open-ended questions, comments and suggestions were analyzed by using content analysis. The same content analysis procedure was used as in Study 1 (see Section 3.1.4 for details).

According to the guideline extraction methodology employed (see Section 3.1.4 for details), no new guidelines were extracted and added to the current set of guidelines. Therefore, Design Guidelines Evaluation 2 consisted of the design guideline set used to develop Prototype B. In addition, according to the participant results, there was no need to make any changes or revisions to the Revised Online Needs Assessment Tool 1. Thus, no second design iteration was conducted and Revised Online Needs Assessment Tool 1 was released as the final tool (Revised Online Needs Assessment Tool 2).

3.2.5 Results

This section presents the analysis of the quantitative and qualitative data gathered from Usability Test Group 2 and 3. In the following segment, the results of the qualitative and quantitative data are presented separately for each questionnaire part.

Part I

Part I consisted of Needs Elicitation Questions in the form of: a) Closed-ended questions with check-box alternatives, b) Closed-ended questions with radio-button alternatives, and c) Open-ended questions (see Section 3.1.2 for more details). The results for each of these different question forms are presented in Appendix S.

Part II

Part II consisted of Usability Questions in the form of: a) Close-ended questions with radio button alternatives and b) Comments and suggestions section (see Section 3.1.2 for more details). The results for each of these question forms are as follows:

a) Close-Ended Questions with radio-button alternatives (9-point Likert-Type scale questions):

The rating distribution chart for the task difficulty levels is shown in Figure 9. As indicated, 93% of the participants provided an overall rating of "7" or above and agreed that the tasks they completed were easy.

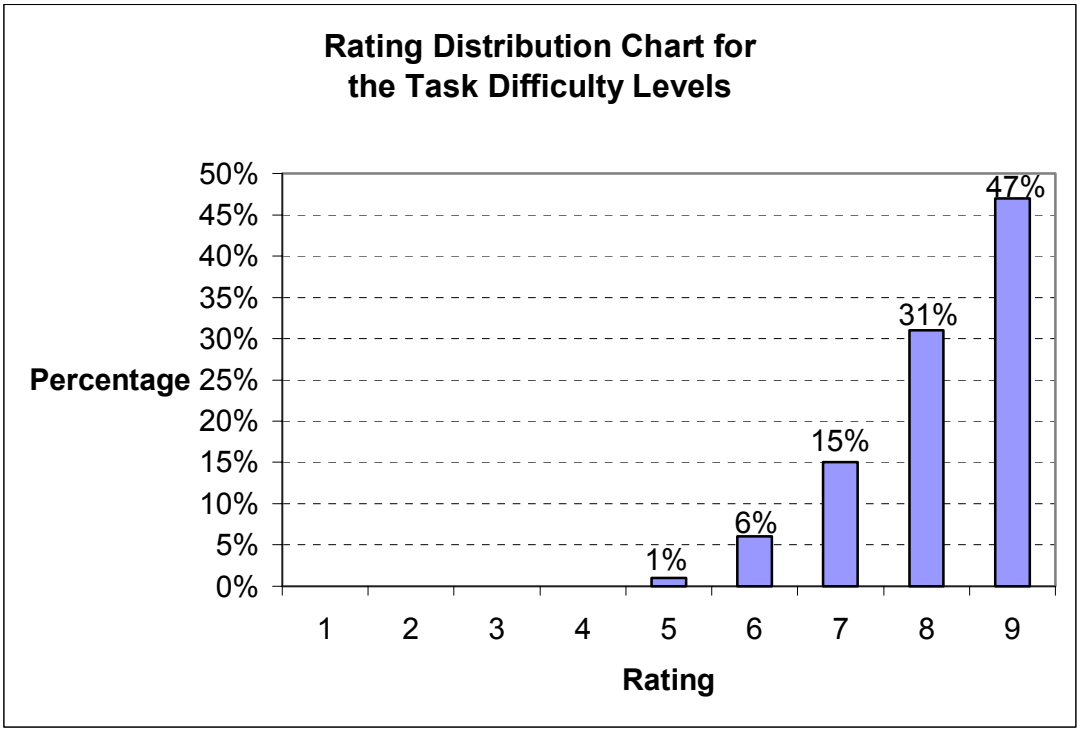


Figure 9: Rating Distribution Chart for the Task Difficulty Levels

The mean ratings and overall rating distribution for each task are provided in Table 27.

Table 27: Task Difficulty

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Rating	SD
Going to the next page	0	0	100	8.50	0.03
Selecting response circles	0	0	100	8.41	0.06
Operating my input devices	0	6.6	93.4	8.25	0.09
Understanding the concept of an "Online Questionnaire"	3.3	6.6	90	8.25	0.11
Selecting responses squares	0	3.3	96.7	8.20	0.05
Opening or getting into the survey	3.3	3.3	93.4	8.03	0.12
Filling in the blanks	0	3.3	96.7	8.00	0.11
Reading the questions	0	13.3	86.7	7.91	0.51
Understanding what is being asked	0	10	90	7.83	0.10

The rating distribution chart for the general usability statements is shown in Figure 10. As indicated, 86% of the participants provided an overall rating of “7” or above and agreed that the Online Tool was easy to use, sufficient, adequate and concise.

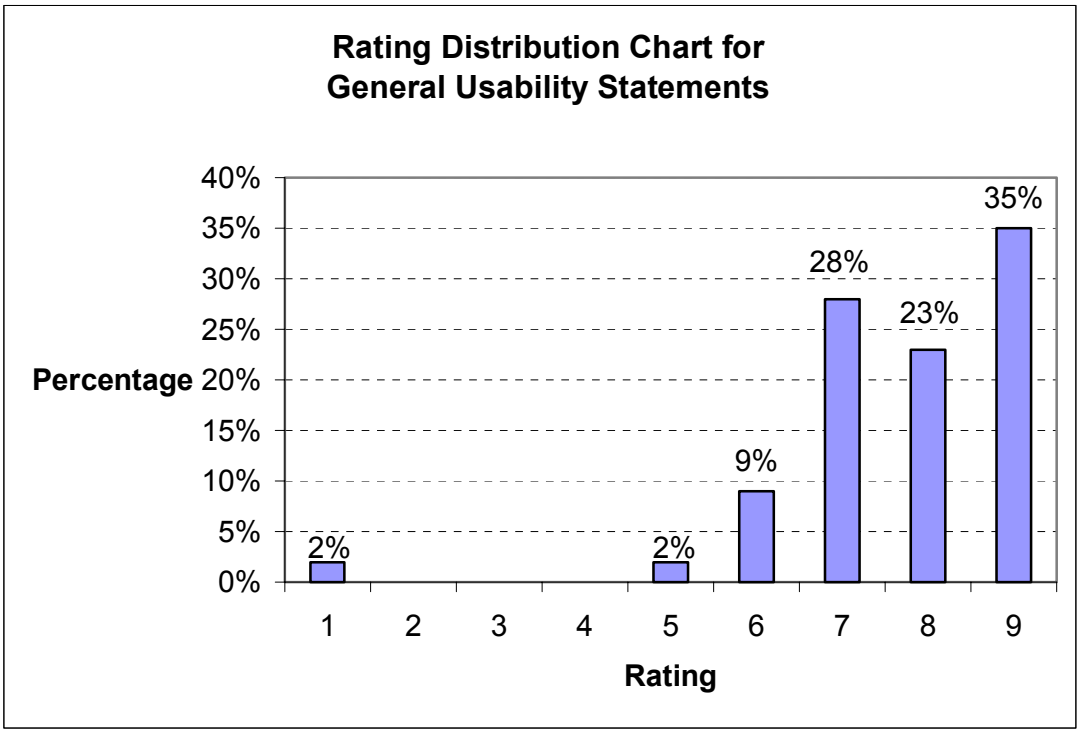


Figure 10: Rating Distribution Charts for General Usability Statements

The mean ratings and overall rating distribution for each statement are provided in Table 22.

Table 22: Usability Statements

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Ratings	SD
I was comfortable with the amount of questions in the Online Questionnaire	**0**	**6.6**	**93.3**	**8.20**	**0.09**
I am comfortable with online evaluations	**0**	**0**	**100**	**8.14**	**0.06**
This Online Questionnaire is easy to use	**3.3**	**3.3**	**93.4**	**8.08**	**0.16**
Overall, the amount of effort it took to complete this questionnaire was acceptable for me	**0**	**3.3**	**96.7**	**8.00**	**0.08**
Overall, the amount of time it took to complete this questionnaire was acceptable for me	**3.3**	**6.6**	**90**	**7.91**	**0.25**
I was comfortable with the content of questions in the Online Questionnaire	**0**	**3.3**	**96.7**	**7.88**	**0.26**
It was easy to view the Online Questionnaire as a link while completing this form	**3.3**	**6.6**	**90**	**7.75**	**0.33**
I would prefer to meet face-to-face with a researcher to evaluate The Online Questionnaire	**3.3**	**16.6**	**63.3**	**5.70**	**0.63**
Using the Online questionnaire helped me communicate what I want and don't want in a Web site	**0**	**13.3**	**86.7**	**7.66**	**0.26**
I was comfortable with the length of the Online Questionnaire	**3.3**	**10**	**86.7**	**7.66**	**0.30**
The questions in the online questionnaire were adequate for me to communicate all my needs and preferences	**3.3**	**13.3**	**83.4**	**7.58**	**0.28**

Subsequent to calculating the mean ratings, guideline extraction methodology was applied to the results of these four questions. However, neither one of the response alternatives met the guideline extraction criteria (see Section 3.1.4 for more details on the criteria). Therefore, no guideline was developed.

In addition to the guideline extraction, 2*2*3 ANOVA was performed for each of the response alternatives completion times. Each ANOVA test was performed for

experience level categories, age and gender (see Appendix R for ANOVA results). The results of these analyses indicated six significantly different findings.

For the task difficulty item “Selecting response squares” (response alternative 4 of question 1), the following significant effect was found:

- There was a significant difference between the mean rating value for female and male participant groups ($F\ (1,12) = 10,\ p < .05$). Tukey’s Studentized Range (HSD) test indicated that the mean rating score for female participants ($M = 8.66,\ SD = 0.65$) was significantly higher than the mean rating score for male participants ($M = 7.83,\ SD = 1.46$). Female participants found selecting response squares significantly easier to use than male participants.
- There was a significant difference between the mean rating value for intermediate and the other two experience level categories ($F\ (2,12) = 11.70,\ p < .05$). Tukey’s Studentized Range (HSD) test indicated that the mean rating value for intermediate experience level ($M = 7.38,\ SD = 1.59$) was significantly lower than the mean rating value for beginner experience level ($M = 8.50,\ SD = 0.755$) and expert experience level ($M = 8.88,\ SD = 0.35$). Intermediate participants found selecting response squares significantly less easier to use than expert and beginner participants.

For the usability statement “It was easy to view the Online Questionnaire as a link while completing this form” (response alternative 2 of question 3, the following significant effect was found:

- There was a significant difference between the mean rating value for female and male participant groups ($F\ (1,12) = 5.33,\ p < .05$). Tukey’s Studentized Range

(HSD) test indicated that the mean rating score for female participants (*M =8.08, SD =0.99*) was significantly higher than the mean rating score for male participants (*M =7.41, SD =1.24*). Female participants found viewing the Online Questionnaire as a link while completing this form significantly easier to view than male participants.

For usability statement “I would prefer to meet face-to-face with a researcher to evaluate The Online Questionnaire” (response alternative 3 of question 3), the following significant effect was found:

- There was a significant difference between the mean rating value for female and male participant groups ($F\ (1,12) = 11.95,\ p < .05$). Tukey’s Studentized Range (HSD) test indicated that the mean rating score for male participants (*M =6.33, SD =1.27*) was significantly higher than the mean rating score for female participants (*M =4.58, SD =2.90*). Male participants expressed a significantly stronger preference towards meeting face-to-face with a researcher to evaluate The Online Questionnaire than female participants.
- There was a significant difference between the mean rating value for older and younger participant groups ($F\ (1,12) = 11.95,\ p < .05$). Tukey’s Studentized Range (HSD) test indicated that the mean rating score for younger participants (*M =6.83, SD =1.80*) was significantly higher than the mean rating score for older participants (*M =4.58, SD =0.77*). Younger participants expressed a significantly stronger preference towards meeting face-to-face with a researcher to evaluate The Online Questionnaire than older participants.

Although the responses to the usability statement “I would prefer to meet face-to-face with a researcher to evaluate The Online Questionnaire” showed participant preference for in-person evaluations, the mean ratings for the other usability statements showed that a more significant overall satisfaction and comfort level with the online questionnaire was achieved.

For completion time, the following significant effect was found:

- There was a significant difference between the mean completion time for older and younger participant groups ($F\ (1,12) = 138.29,\ p < .05$). Tukey’s Studentized Range (HSD) test indicated that the mean completion time for older participants ($M = 8.25,\ SD = 0.97$) was significantly higher than the mean completion time for younger participants ($M = 4.58,\ SD = 0.78$).

The Cronbach coefficient alpha value was *.73* and the range of item-to-total correlation was *.03 – .74*. The correlation coefficients among the 20 response alternatives were all positive and statistically significant at the *.05* level. The correlation coefficients ranged from a low *.03* to a high *.74*, indicating the degree of the relationships between the response alternatives. The lowest reliable variance of the total scores was at least 73%. Each response alternative had adequate internal consistency and there was no need to further improve the reliability of these questions and their response alternatives.

<u>b) Comments and suggestions sections:</u>

In this section, participants were asked to provide feedback regarding the questionnaire, participation process, needs, requirements and preferences. For the results gathered from this section, content analysis was performed. Comment was

identified as final response category (as shown in Tables 26 and 17) and guideline extraction methodology was applied.

Table 26: Qualitative Responses (Likes)

	Comments	Frequency
Younger Users	Questionnaire was short, concise and straightforward	1
	Questionnaire asked all the right questions	1
	Questionnaire was complete	3
	Questionnaire covered everything	4
Older Users	Questionnaire was concise and short	4
	Questionnaire was easy	4
	Questionnaire was straightforward and precise	5
	Questionnaire was listed in a very easy to follow pattern	2
	So it is nice to do a survey where your opinion is asked and counts. Healthcare has become so untouchable for most of us, we just have to do what the insurance companies tell us to do.	1

Table 17: Qualitative Responses (Dislikes)

	Comments	Frequency
Younger Users	No dislikes	8
	Would be nice if it was more colorful and had some animation. But I don’t see how you could do that for a simple survey	1
Older Users	I didn't like or found to be difficult or offensive.	2
	No dislikes	6
	The set up was very good however the placement of choices are usually (on most surveys) right beside the question would be 1 -Excellent 2 - then 9 would have been the Least effective means not vice versa.	2

The responses for comments category consisted of phrases participants used to communicate their satisfaction with the tool. One of the comments included an observation regarding an alternative usability scale design. However, this design alternative did not exceed the overall score 0.4 (see Section 3.1.4 for more details on calculating the score). Therefore, no new design guideline was extracted for these two response categories.

Part III

Part III consisted of Guideline Evaluation Questions in the form of: a) Close-ended questions with radio button alternatives and b) Comments and suggestions section (see Section 3.1.2 for more details). The results for each of these question forms are as follows:

<u>a) Close-Ended Questions with radio-button alternatives (9-point Likert-Type scale questions):</u>

The Rating Distribution Chart for Agreement Scale is shown in Figure 11. As indicated, 83% of the participants provided an overall rating of "7" or above and agreed that each design guideline of Prototype B was well implemented.

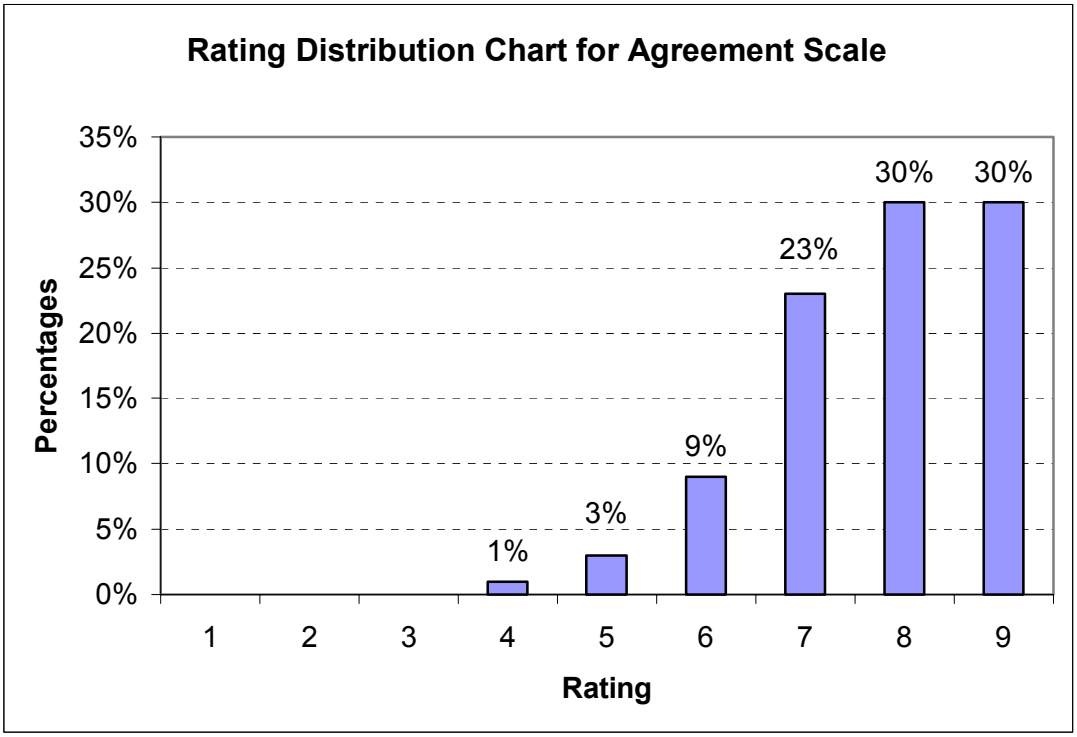


Figure 11: Rating Distribution Chart for Agreement Scale

Agreement scale mean ratings and overall rating distribution for each guideline are provided in Table 16.

Table 16: Agreement Scale Mean Ratings and Overall Rating Distribution

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Rating	SD
Provide easy to use help during office and non-business hours	**0**	**3.3**	**96.7**	**8.08**	**0.03**
Color is properly used for the characters, foreground, background, and to highlight important information	**0**	**6.6**	**93.4**	**8.00**	**0.06**
The Online Questionnaire works properly on your computer and your browser	**0**	**3.3**	**96.7**	**8.00**	**0.06**
The information presented is relevant	**0**	**6.6**	**93.4**	**7.91**	**0.15**
Links, buttons, keys and icons are sufficient, easily distinguished from other text and easy to access	**3.3**	**6.6**	**90**	**7.91**	**0.25**
Information presentation and operating procedures are logical and consistent with other commonly used applications and throughout the Online Questionnaire	**0**	**3.3**	**97.6**	**7.91**	**0.16**
Visual field is placed in narrow view, the use of open space is balanced and sufficient right margin is left for various window/display sizes	**0**	**3.3**	**96.7**	**7.91**	**0.13**
Graphics, distracting background elements and animation are kept to a minimum or voided altogether	**3.3**	**3.3**	**93.4**	**7.79**	**0.16**
Font, typeface (for example, bold, italic, underline, capital letters etc.), and size are properly applied for the text	**3.3**	**10**	**86.7**	**7.66**	**0.23**
Scrolling to view information is minimized. Horizontal is avoided	**10**	**6.6**	**83.4**	**7.66**	**0.25**
Page layout (placement or spacing of headings, types, paragraphs, labels or other page elements) is properly designed	**6.6**	**13.3**	**80**	**7.63**	**0.25**
Links, buttons, keys and icons are of proper size and color, clearly labeled, and meaningful	**6.6**	**10**	**80**	**7.63**	**0.26**
Information layout is not cluttered. Language is clear, simple and not technical	**3.3**	**10**	**86.7**	**7.54**	**0.33**
Open fields, check boxes, and radio buttons are sufficient. Typing demands are reduced	**10**	**10**	**80**	**7.38**	**0.51**

The Rating Distribution Chart for Importance Scale is show in Figure 12. Ninety-Two percent (92%) of the participants provided an overall rating of “7” or above and agreed that each design guideline of Prototype B was important.

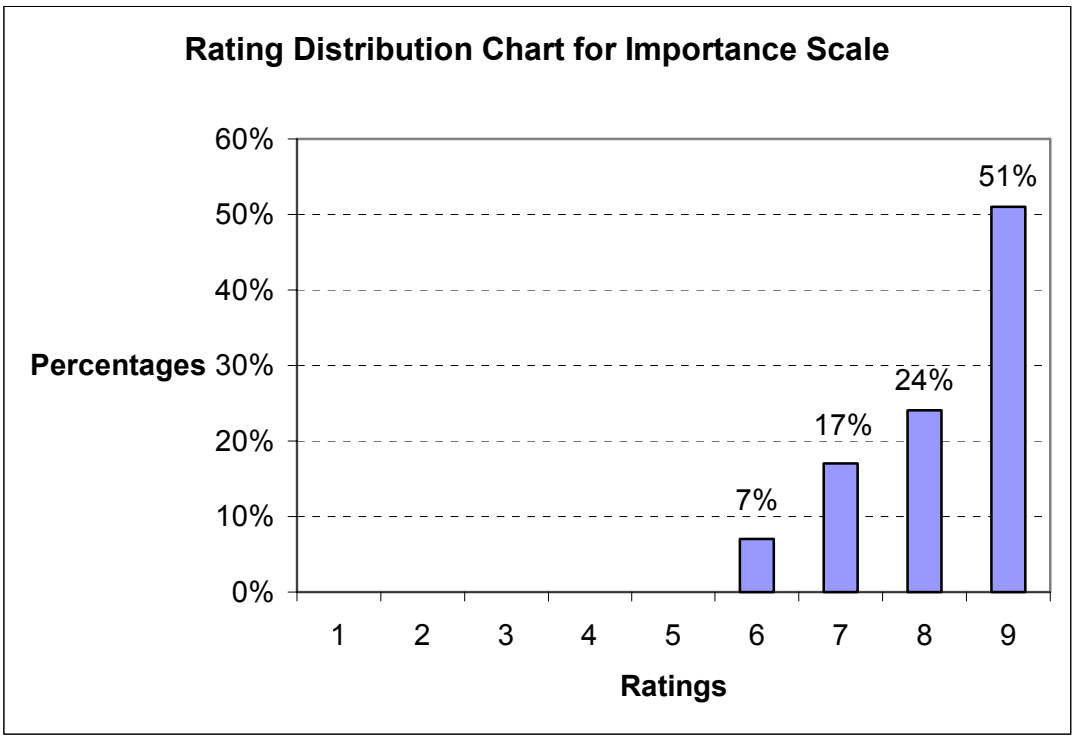


Figure 12: Rating Distribution Chart for Importance Scale

Importance scale mean ratings and overall rating distribution for each guideline are provided in Table 12.

Table 12: Importance Scale Mean Ratings and Overall Rating Distribution

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Rating	SD
Links, buttons, keys and icons are of proper size and color, clearly labeled, and meaningful	**0**	**0**	**100**	**8.54**	**0.03**
Page layout (placement or spacing of headings, types, paragraphs, labels or other page elements) is properly designed	**0**	**0**	**100**	**8.41**	**0.04**
Links, buttons, keys and icons are sufficient, easily distinguished from other text and easy to access	**0**	**3.3**	**96.7**	**8.25**	**0.06**

Item	% ratings that were 3 <… ≤ 5	% ratings that were 5 <… ≤ 7	% ratings that were 7≤	Mean Ratings	SD
Graphics, distracting background elements and animation are kept to a minimum or voided altogether	**0**	**6.6**	**93.4**	**8.25**	**0.10**
Open fields, check boxes, and radio buttons are sufficient. Typing demands are reduced	**0**	**6.6**	**93.4**	**8.25**	**0.10**
Visual field is placed in narrow view, the use of open space is balanced and sufficient right margin is left for various window/display sizes	**0**	**6.6**	**93.4**	**8.20**	**0.13**
The Online Questionnaire works properly on your computer and your browser	**0**	**3.3**	**96.7**	**8.20**	**0.06**
Provide easy to use help during office and non-business hours	**0**	**0**	**100**	**8.20**	**0.04**
Color is properly used for the characters, foreground, background, and to highlight important information	**0**	**16.6**	**83.4**	**8.12**	**0.15**
Information layout is not cluttered. Language is clear, simple and not technical	**0**	**6.6**	**93.4**	**8.08**	**0.15**
Scrolling to view information is minimized. Horizontal is avoided	**3.3**	**3.3**	**93.4**	**8.00**	**0.26**
Information presentation and operating procedures are logical and consistent with other commonly used applications and throughout the Online Questionnaire	**0**	**6.6**	**93.4**	**8.00**	**0.23**
Font, typeface (for example, bold, italic, underline, capital letters etc.), and size are properly applied for the text	**0**	**10**	**86.7**	**7.83**	**0.51**
The information presented is relevant	**0**	**10**	**86.7**	**7.70**	**0.98**

Subsequent to calculating the mean ratings, the guideline extraction methodology was applied to the results of these fourteen guidelines. However, neither one of the response alternatives met the guideline extraction criteria (see Section 3.1.4 for more details on the criteria). Therefore, no guideline was developed.

In addition to the guideline extraction, 2*2*3 ANOVA was performed for each of the response alternatives and for completion times. Each ANOVA test was performed for experience level categories, age and gender (see Appendix R for ANOVA results).

The results of these analyses indicated that no guideline questions showed significantly different results. However, no significant effects were found

The Cronbach coefficient alpha value was .85 and the range of item-to-total correlation was *.04* – .63. The correlation coefficients among the 20 response alternatives were all positive and statistically significant at the *.05* level. The correlation coefficients ranged from a low *.04* to a high *.63*, indicating the degree of the relationships between the response alternatives. The lowest reliable variance of the total scores was at least 85%. Each response alternative had adequate internal consistency and there was no need to further improve the reliability of these questions and their response alternatives.

b) Comments and suggestions sections:

In this section, participants were asked to provide any other guideline that they thought was important and should be included in the design of the Online Questionnaire. For the results gathered from this section, content analysis was performed. Comments were identified as final response category (as shown in Table 9) and guideline extraction methodology was applied.

Table 9: Qualitative Responses

Responses		Frequency
Younger Users	It was short and to the point.	3
	I don't think I have any concerns. The guidelines were sufficient and covered everything	5

Responses		Frequency
Older Users	Liked it overall	5
	This is the first survey of this type that I have done online. I have not had any major problems so far, but I really do not know much about these guidelines	2
	I found the guidelines and instructions to be simple to follow and the questions were thought evoking.	4
	I didn't find any problems with the guidelines that were set down before the survey was started and after the survey began.	3

The responses for comments category consisted of phrases participants used to communicate their satisfaction with the guidelines and the online tool. Thus, no new design guideline was extracted for this response category. Since no new guidelines were extracted according to the results of Study 2, the Design Guidelines Evaluation 2 will consist of the design guideline set (see Section 3.1.3 for the guideline set).

3.3 Discussion

The discussion of the findings of this research is presented in four sections. (1) A summary of the results presented in the previous sections. (2) Further observations and interpretations of the findings. (3) Recommendations for future remote usability and needs assessment studies. (4) Conclusions.

3.3.1 Summary of Results

This thesis project was designed to explore the concept of Adult Age-Independent ONATs and Adult Age-Independent Usability Design Guidelines. The primary research objective of this project was to address the question: To what degree can one ONAT be developed to be usable for both older and younger users? This project used two different studies to develop a usable Adult Age-Independent ONAT. During these studies, two different design iterations were conducted and the Revised Online Needs Assessment Tool 2 (see Section 3 for details) was developed. To examine the extent to which this tool was usable for both older and younger adults, participants were asked to provide ratings and comments in the areas of navigation and selection tools, information design and layout, satisfaction, ease of use, comprehension, content, design guidelines and other relevant tool features. These results showed that complete user satisfaction was achieved in all these areas for the Adult Age-Independent ONAT.

In Study 1, 88% of the participants agreed that the Online Tool was easy to use, effective, adequate and concise, 84% of the participants agreed that each design guideline was successfully implemented and 93% of the participants agreed that each design guideline was important by providing an overall rating of “7” or above.

In Study 2, 86% of the participants agreed that the Online Tool was easy to use, effective, adequate and concise, 83% of the participants agreed that each design guideline was successfully implemented and 92% of the participants agreed that each design guideline was important by providing an overall rating of “7” or above. For both studies, while several participants have communicated their overall satisfaction with the length, content and adequacy of the tool, no participant communicated any dissatisfaction, concern or problem.

The secondary research objective of this project was to develop adult age-independent usability guidelines for online applications with healthcare information content. During this project, several guidelines were extracted, tested and refined. As a result, the second research objective was achieved by developing the final usability guideline set, Design Guideline Evaluation 2. In addition to achieving the two research objectives, the usability test results showed that:

- The completion times for older users were significantly higher than the completion times for younger users for Part I and II. Even though the completion time for Part III was not statistically significant, the mean for older adult users was higher than the mean for younger adult users. This finding was also supported by the results of several prior studies on the differences in information processing and reaction time between older and younger users (e.g., Czaja, 1997; Vercruyssen, Carlton and Diggles-Buckles, 1989).
- All the questions had adequate internal consistency and there was no need to improve reliability.

- There were no significant correlation between the rating scores and the experience levels.
- There were no significant trends for the independent variables. Although female participant ratings were higher than male participant ratings for most of the questions, no statistical significance was found.
- Open-ended questions failed to elicit sufficient amount of qualitative information. It was seen that older users provided more information provided for open-ended questions than younger users. However, the total amount of qualitative results gathered from older participants was insufficient for an in-depth content analysis.
- There were no additional guidelines extracted from the results of Study 1 and 2. All the usability guidelines were also rated to be important and successfully implemented. Therefore, Design Guideline Evaluation 1 & 2 consisted of the initial usability guidelines (see Section 3.1.3 for a complete list) gathered during the preliminary development activities.

3.3.2 Interpretation of Findings

In addition to exploring the concept of Adult ONAT and Adult Age-Independent Usability Guidelines, the results of this study identified pertinent issues and areas of discussion. The following section includes a detailed discussion of each issue that is recommended for further investigation:

a) Usefulness of open-ended questions on online questionnaires:

This research used open-ended questions as the only source of questionnaire qualitative data. Several factors were identified as the possible underlying cause(s) for the insufficient amount of qualitative data elicited.

One of the reasons why participants chose not to provide sufficient information as an answer to the open-ended questions could be practicability. The participants might have lacked the information or the time to provide a response at the time this tool was being completed. In which case, providing the means of reporting additional responses to the open-ended questions at a later time could be useful.

The effectiveness of using open-ended questions in gathering usability information should be investigated. A future study could be used to determine the effectiveness of only using closed-ended questions with response alternatives and ratings to collect usability information. This investigation should also concentrate on providing a complete set of questions, which would have broad content and application area to include all aspects of usability evaluation at hand.

The effects of conducting remote versus laboratory-based studies should also be considered. Various studies have reported that remote users are found to be less motivated to report positive feedback compared to laboratory-based users (Capra,

2001; Thompson, 1999). These studies have suggested that the laboratory environment and the interaction with the experimenter could motivate the participants to act prejudiced. It has also been proposed that in laboratory-based studies, participants frequently tend to please the experimenter by praising the tool they are evaluating. However, as remote studies lack the presence of the experimenter and his/her potential impact, the elimination of this false appraisal could result in the amount of qualitative responses.

The effects of participant motivation should also be considered as a possible cause in the amount of the qualitative information gathered. Various studies have suggested alternative ways of motivating participants to report more feedback in remote usability evaluations. Thompson (1999) has suggested conducting the reporting section of remote usability evaluations as a collaborative effort among a group of users, who interact with the same interface. Besides the motivational benefits of this joint evaluation, the effects of group-think and social conformity on the quality and the originality of the data needs to be addressed by any empirical data.

b) Reliability of Guideline Extraction Criteria:

Due to the absence of guideline extraction criteria in the existing literature, a new guideline extraction methodology was created and used for this research. This guideline extraction methodology was not tested to determine whether or not it was a reliable measure that consisted of appropriate extraction criteria. With the exception of one content item, the extraction methodology was found to be appropriate for this research. The results showed that two participants during both

Study 1 and 2 identified “information on HIV and other sexually transmitted diseases” as a content item they would like to have on a healthcare web site. However, this content item failed to meet the set guideline extraction criteria and therefore, it was not included in the next design iteration. For this research, exclusion of this content item from the next design iteration did not show any significant effect on the findings or the overall design. However, in order to avoid the exclusion of important items or guidelines that might have significant effects on the results of future studies, the extraction methodology should be tested for repeating responses for consecutive studies. The extraction methodology should be tested for other remote usability and needs assessment applications to further investigate the efficacy and limitations of the tool. Further research should also explore the reliability and usefulness of other alternative methods with which qualitative data results can be quantified.

c) <u>The need for summative evaluation/further research</u>:

Due to the budget and time restrictions, this study consisted of only formative evaluation. In order to release the online tool and the guidelines as products, summative evaluation needs to be performed. Besides addressing the previous two areas of discussion, summative evaluation or future research should:

a) be conducted with a larger, more diverse group of participants. This research was conducted in Blacksburg, Virginia where the community mainly consisted of Virginia Tech affiliates. In order to accommodate users with a variety of cultural and educational backgrounds, the participant diversity has to be achieved.

b) revaluate participant eligibility measures. This study used “visiting 2-3 web sites prior to the study” as a qualification requirement. Before recruiting every participant, they were asked to confirm that they fulfilled this requirement. However, during Study 1, one participant mentioned that he/she was not able to understand the concept of a healthcare web site. The participant’s data was examined and the validity of her responses was ensured. As a result, a general definition of this term was included during the next design iteration. This change was not found to have any significant effect on the results or the overall design of the tool. However, it showed a need for using an alternative method such as a short task or a questionnaire to confirm participant eligibility.

c) be an on going and constantly improving activity. As every technical industry, healthcare profession is fast paced and constantly changing. This constant change will demand a continuous update of the questionnaire format, features and contents. In order to achieve this constant improvement, the questionnaire should be administered periodically and revised according to the results.

d) explore gender differences on rating scores. Even though it was not a significant trend, the study results showed that female participant ratings were higher than male participant ratings for most of the questions. Due to the limited scope of the study, the additional data were not collected to further examine this finding. However, this finding should be examined during the summative evaluation.

e) explore special user needs. People with disabilities and special needs are other areas summative evaluation should examine. Today, most of the Internet and W.W.W. applications offer full access to users with physical and sensory disabilities. These applications are found to not only aid the growing population of people with disabilities and older users but they are also found to be very beneficial to people from different adult age groups. Further research in this field will enable the online questionnaire to be more accessible and beneficial to all adult age groups.

f) investigate the use of alternative interface modalities. The Adult ONAT was primarily developed as text-based user interface. Previous studies have concentrated on the use of auditory or visual clues, peripheral tasks and their effects on mental workload and interruption of concentration (McDowd & Birren, 1990; Sit & Fisk, 1997; Zandri & Charness, 1989). However, no studies have expanded the research to other interface modalities and researched the effectiveness of and usability issues related to using narrated video clips, haptic interfaces and virtual reality for all adult age groups. These interface modalities could be developed to provide user-centric portrayal of online needs assessment and usability evaluation tools that could be applied to other remote assessment techniques and Internet applications for all adult age groups.

3.3.3 Recommendations

Throughout this project, the following recommendations were found to be very helpful in making the tool development process faster, decreasing the amount of technical difficulties and achieving high response rate and useful data:

- Verify participant eligibility. Do not rely on participant opinion.
- Provide the URL of the survey in a written format as well as an electronic format. Also include explanations for alternative ways of accessing the survey such as copying and pasting.
- Design the surveys compatible with older computer processors, older software versions and slower Internet connections.
- Pretest the surveys with the oldest computer processors, the oldest software versions, and the slowest Internet connections that are in current use.
- Select the programming language and the software(s) according to server platform you are going to use. This will avoid compatibility problems while uploading surveys on the web.
- Select the software(s) and the programming language(s) according to the survey format and features. This will save time during the survey development process and avoid browser compatibility problems.
- Design the surveys for the least knowledgeable and the least experienced users.
- Design short, precise and easy to use surveys. If the survey has to be long, break the survey into sections. This will avoid data loss for all sections in case of a technical problem. Also, do not make any surveys that take longer than 45 minutes to complete.

- Control and protect the survey data by assigning usernames and passwords to your participants. If it is not necessary to block access by creating login protection, assign distinctive usernames and passwords.
- Be accessible to participants in case they have a problem or a question during the course of data collection. Be sure to provide participants a phone number and an email address where project director can be reached.
- Supply a hardcopy of every electronic help and reminder document provided.
- Know the community from which participants are recruited. Be knowledgeable about the common working/business hours, leisure hours and vacation/traveling seasons for all the age groups of the target community. This will help with participant recruitment in areas where there is a limited participant resource.
- Prefer to use more personal contact methods with older users. Chose in-person contact and telephone call over email. Be prepared to make presentations to local groups instead of posting flyers.
- Be prepared to distinguish between the problems reported due to software/hardware malfunction issues on the user and the experimenter side. Keep in mind that some of the problems reported would be related to user equipment and/or software even though they may be reported otherwise.
- Be resourceful about the most commonly used Internet Service Providers in the targeted area. As some compatibility and other technical issues are likely to arise, the project director needs to know how to troubleshoot or seek resolution.
- Avoid using software packages that automatically generate the programming code for the survey. If one of these software packages is used, select the ones

that do not use unfamiliar or complex programming languages. This will eliminate future complications if changes need to be made.

3.3.2 Conclusions

This project was successful in developing and evaluating a final Adult Age-Independent ONAT (Revised Online Needs Assessment Tool 2) and Adult Age-Independent Design Guidelines (Design Guideline Evaluation 2).

Revised Online Needs Assessment Tool 2 consisted of three sections that were used to conduct remote needs assessment. These sections allowed the participants to provide user requirement information, ratings of overall usability and design guideline evaluation. The results of Study 1 & 2 showed that the tool achieved a high level of overall user satisfaction for accommodating user requirements for older and younger adults. Design Guideline Evaluation 2 consisted of adult age-independent usability guidelines for online applications with healthcare information content. Preliminary development activities assured that all usability guidelines that are applicable to ONAT were compiled. The results of Study 1 & 2 showed that the guideline set was found to be comprehensive, important and successfully implemented by older and younger adults.

In addition to the tools developed, the results of this research provided usability specialists' and interface designers' a knowledge resource for older and younger adult user healthcare information needs and usability information about needs assessment tools, design guidelines, features and formats supported. The results also demonstrated remote needs assessment and usability evaluation as more feasible alternatives to laboratory-based studies for older and younger adults. Remote needs assessment,

which provides a cost-effective alternative by addressing situations where users can not be brought into a laboratory environment due to mobility or location problems, was found to be comfortable, sufficient and satisfactory by the users.

References

Adams, A. S., Thieben, K. A. (1991). Automatic teller machines and older population. *Applied ergonomics,* 22, 85-90

Brickfield, C. F. (1984). Attitudes and perceptions of older people towards technology. In P. K. Robinson and J. E. Birren (Eds.), *Aging and Technological Advances* (pp. 31-38). New York: Plenum Press.

Brouwer-Janse, M. D., Suri, J. F., Yawitz, M. A., Vries, G. D., Fozard, J. L., Coleman, R. (1997). User interfaces for young and old. *Interactions*, 4(2), 34-46

Capra, G. M. (2001). *An Exploration of End-User Critical Incident Classification.* Unpublished thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA

Charness, N., Urchman, C. E., Boritz, G. A. (1992). Training older adults in word processing: Effects of age, training technique and computer anxiety. *International journal of aging and technology,* 5, 79-106

Cooper, B. P., Lee, M. D., Goska, R. E., Anderson, M. M., Gay, P. E. Jr., Fickes, A., Fisk, A. D. (1992). Age-Related Effects in Consisted Memory Search: Performance is the Same but What about Learning? *In Proceedings of the Human Factors and Ergonomics Society 36th Annual Meeting* (pp. 146-150). Santa Monica, CA: Human Factors and Ergonomics Society.

Costello, R. B. et al. (1993). *The American Heritage College Dictionary* (3rd edition). Boston, USA: Houghton Mifflin Company.

Czaja, S. J. (1997). Computer Technology and the Older Adult. In Helander M., Landauer T. K. and Prabhu P. (Eds.), *Handbook of Human-Computer Interaction* (pp. 797-813). Amsterdam, The Netherlands: Elsevier Science B. V.

Czaja, S. J., Clark, M. C., Weber, R. A., Nachbar, D. (1990). Computer Communication Amongst Older Adults. *In Proceedings of the Human Factors and Ergonomics Society 34th Annual Meeting* (pp. 146-148). Santa Monica, CA: Human Factors and Ergonomics Society.

Danowski, J. A., Sacks, W. (1980). Computer communication and elderly. *Experimental aging research,* 6,125-135

Ellis, R., Kurniawan, D., Sri, H. (2000). Increasing the usability of online information for older users: A case study in participatory design. *International Journal of Human-Computer Interaction,* 12(2), 263-276

Feinberg, S., Johnson, G., Peter, Y. (1997). Designing and testing customer satisfaction surveys on WWW sites. *Proceedings/STC, Society for Technical Communication Annual Conference 1997. Soc for Technical Communication,* 298-300

Gable, R. K., Wolf, M. B. (1993). Instrument development in the affective domain (2nd edition). Norwell, MA: Kluver Academic Publishers

Garner, R. (2000). Search Engine Marketing. Retrieved June 30, 2005, from Google.com database [online]. Available: http://rgarner.com/usability.html

Garson, D. G. (2004). *Scales and Standard Measures.* North Carolina State University Social Science Research and Computing Laboratory, Raleigh, NC, from Google.com database [online]. Available: http://www2.chass.ncsu.edu/garson/pa765/standard.htm

Haigh, R. (1993). The ageing process: A challenge for design. *Applied Ergonomics,* 24(1), 9-14

Hartson, H. R., Castillo, J. C., Kelso, J., Neale, W. C., Kamler, J. (1996). Remote evaluation: the network as an extension of the usability laboratory. *In Proceedings of the Conference on Human Factors in Computing Systems (CHI 96)*, 228-235. New York, USA: ACM Press

Hawthorn, D. (2000). Possible implications of aging for interface designers. *Interacting with Computers,* 12(5), 507-528

Health On The Net Foundation (1999, April). Health On The Net Foundation's Fourth Survey on the Use of the Internet for Medical and Healthcare Purposes. Retrieved August 1, 2001, from Excite.com database [online]. Available: http://www.hon.ch/Survey/ResumeApr99.html

Humprey, D. G., Kramer, A. F., Gore, S. S. (1994). Perceptual organization and grouping factors: Age related effects. *In Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting* (pp. 170-174). Santa Monica, CA: Human Factors and Ergonomics Society.

McDowd, J. M., Birren, J. E. (1990). Aging and attentional processes. In Birren J. E., Schaie K. W. (Eds), *Handbook of the psychology of aging* (222-233). New York, USA: Academic.

Moravec, H. P. (1999). Power and Presence. In Moravec H. P. (Ed.), *Robot: Mere Machine to Transcendent Mind* (pp. 51-91). New York, USA: Oxford University Press, Inc.

Morrell, R. W. (2000). A Survey of World Wide Web Use in Middle-Aged and Older Adults. *Human Factors*, 42(2), 175-182

Nielsen, J. (1993a). Categories of Users and Individual User Differences. In Nielsen J. (Eds.), *Usability Engineering* (43-48). MA, USA: Academic Press, Inc.

Nielsen, J. (1993b). The Usability Engineering Life Cycle. In Nielsen J. (Eds.), *Usability Engineering* (71-115). MA, USA: Academic Press, Inc.

Nielsen, J., Molich, R. (1990). Heuristic Evaluation of user interfaces. In Chew, J. C. and Whiteside, J. (Eds.), *Empowering people* (pp 249-256), [CHI '90 Conference Proceedings] New York, USA: ACM Press.

Nielsen, J., Schaefer, L. (1993). Sound effects as an interface element for older users. *Behaviour & Information Technology,* 12(4), 208-215

Norman, D. A. (1990). User-Centered Design. In Norman D. A. (Ed.), *The Design of Everyday Things* (pp. 187-219). New York, USA: Doubleday/Currency Dell Publishing Group, Inc.

Preece, J. (1993). A Guide To Usability: Human Factors in Computing. In Preece J., Benyon D., Davies G., Keller L., Rogers Y. (Eds.), *Evaluation* (pp. 107-120). Great Britain: Addison-Wesley/Henry Ling Ltd.

Reaves, C. C. (1992). Statistical Decisions. In Swain E. (Ed.), *Quantitative Research for Behavioral Sciences* (236-240). Canada: John Wiley & Sons, Inc.

Reviere, R., Berkowitz, S., Carter, C. C., Ferguson, C. G. (1996). Needs Assessment: A Creative and Practical Guide for Social Scientist. In Reviere, R., Berkowitz, S., Carter, C. C., Ferguson, C. G. (Eds), *Introduction: Setting the Stage (pp. 1-11)*. London, UK: Taylor and Francis ltd.

Rogers, A. W., Fisk, D. A., Mead, E. S., Walker N. (1996). Training older adults to use automatic teller machines. *Human Factor,* 38(3), 425-433

Rogers, W. A., Gilbert, D. K., Cabrera, E. F. (1994). An in-depth analysis of automatic teller machine usage by older users. *In Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting* (pp. 142-146). Santa Monica, CA: Human Factors and Ergonomics Society.

Sharit, J., Czaja, S. J. (1994). Ageing, computer-based task performance, and stress: issues and challenges. *Ergonomics,* 37(4), 559- 577

Shaw, R. J. (1990). Older adults sometimes benefit from environmental support: Evidence from reading distorted text. *In Proceedings of the Human Factors and Ergonomics Society 34th Annual Meeting* (pp.128-132). Santa Monica, CA: Human Factors and Ergonomics Society.

Sit, R. A., Fisk A. D. (1996). Retention of multiple-task performance: Age-related differences. *In Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting* (pp.128-132). Santa Monica, CA: Human Factors and Ergonomics Society.

Smith-Jackson, T. L., Williges, R. C. (2000). The Role of Needs Assessment and Remote Usability Methodology in Designing Telesuport Systems for Users with Disabilities. In Winters J. (Eds.), *Proceedings of the Resna 2000 Annual Conference* (564-569). Arlington, VA, USA: Resna Press.

Smith-Jackson, T. L., Williges, R. C., Kwahk, J., Capra, M., Durak, T. (2001). *User requirements specification for a prototype healthcare information Web site and an online assessment tool*. Technical Report # ACE/HCIL-2001-01-01. Virginia Polytechnic Institute and State University, Grado Department of Industrial and Systems Engineering.

Solomon, M. (1993). Content Analysis: A Potential Tool in the Searcher's Arsenal. *Database*, v16 n2 (62-68)

Sutherland, S (1995). *The Macmillan Dictionary of Psychology* (2nd edition). London, UK: The Macmillan Press, Ltd.

The associated press (1999, May 17). Computer companies cater to elderly. In the freedomforum.org (a nonpartisan, international foundation dedicated to free press, free speech and free spirit for all people) [online]. Available: http://www.freedomforum.org/templates/document.asp?documentID=11545

Thompson, J. A. (1999). *Investigating the effectiveness of applying the Critical Incident Reporting Technique to Remote Usability Evaluation.* Unpublished thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA

U. S. Census Bureau (1993, October). Computer use in the United States. In U. S. Census Bureau website [online]. Available: http://www.census.gov/population/socdemo/computer/report93/comp1.txt

U. S. Census Bureau (1997, October). Purpose of Internet Use at Home by People 18 years and over. In U. S. Census Bureau website [online]. Available: http://www.census.gov/population/socdemo/computer/report97/tab14.pdf

Vercruyssen, M., Carlton, B. L., Diggles-Buckles, V. (1989). Aging, reaction time, and stages of information processing. *In Proceedings of the Human Factors and Ergonomics Society 33rd Annual Meeting* (pp.174-178). Santa Monica, CA: Human Factors and Ergonomics Society.

Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, 34(4), (457-468)

Weisbord, M. R. (1987). Lewin: The Practical Theorist. In Weisbord, M. R. et al. (Eds.), *Productive workplaces: Organizing and managing for dignity, meaning, and community* (pp. 70-88). San Francisco, USA: Josey-Bass Publishers

White, H., McConnell, E., Clipp, E., Bynum, L., Teague, C., Navas, L., Craven, S., Halbrecht, H. (1999). Surfing the Net in Later Life: A Review of the Literature and Pilot Study of Computer Use and Quality of Life. The Journal of Applied Gerontology*, 18(3), 358-378*

Williges, R. C., Smith-Jackson, T. L., Kwahk, J. (2001). *User-Centered Design of Telemedical Support Systems for Seniors*. Technical Report # ACE/HCIL-2001-01-02. Virginia Polytechnic Institute and State University, Grado Department of Industrial and Systems Engineering.

Wilson, J. R., Corlett, E. N. (1990). Basic ergonomics methods and techniques. In Wilson, J. R., Corlett, E. N. (Eds.), *Evaluation of Human Work* (pp. 141-229). Pa, USA: Taylor & Francis Ltd.

Worden, A., Walker, N., Bharat, K., Hudson, S. (1997). Making computers easier for older adults to use: area cursors and sticky icons. *In Conference proceedings on Human factors in computing systems 1997* [online]. Available: http://www1.acm.org/sigs/sigchi/chi97/proceedings/paper/

Zandri, E., Charness, N. (1989). Training older and younger adults to use software. Special issue: cognitive aging: Issues in research and application. *Educational Gerontology,* 15, 615-631

APPENDIX A

THE GVU GUIDELINES

Feinberg and Johnson (1997) stated the design guidelines developed from the empirical studies of Graphics, Visualization, and Usability Center of the Georgia Institute of Technology as:

1) Place the most important questionnaire at the top of the page (users get tired during surveys).
2) Place the questionnaire with the most questions near the top of the page (users read questionnaires placed first for longer periods of time).
3) Understand the trade-off between gathering sensitive information and attrition (users are sensitive to privacy issues).
4) Enforcing question completion does not drastically increase attrition ("question-answer-adapt/re-ask" cycle does not necessarily drive users away). (p. 298)

APPENDIX B

GUIDELINES PROPOSED BY ELLIS, KURNiAWAN (2000)

The summary of the design recommendations developed from participatory design process regarding the format and display by Ellis and Kurniawan (2000) is as follows:

1) Use a sans serif font, such as Arial or Helvetica. The literature generally supports the contention that older users perform better with these fonts, and our users found them more appealing and subjectively easier to read than the serif font (New Times Roman) that was present in the original design. The observation is consistent with the findings of Morrell and Echt (1997).
2) Use dark type on a light background (commonly referred to as “negative contrast” or positive image”), emphasizing high contrast. We make this recommendation based on our experience with this group of users, despite the fact that the literature is mixed with regard to the positive-negative contrast issue (Murch, 1987; Snyder, 1988; Tobias, 1987)
3) Make sure that links (a) are placed where they are easy to see, (b) are fairly large (the size settled on by the PD team here was 180 pixels * 22 pixels for a graphic button), and (c) have plenty of dead space around them to prevent accidental selection. This confirms and extends the suggestions of Czaja (1997) and Kelley and Charness (1995).
4) Plan screen sizes to fit in lower resolution display setting, at least 800*600 and perhaps even 640*480 for monitors up to a 17-inch screen size. Plan for users to browse with a single maximized (full-screen) window, and minimize the use of links and references to new browser windows. (p. 273-274)

APPENDIX C

GUIDELINES PROPOSED BY Smith-Jackson et al. (2001)

A brief summary of the design recommendations developed from the previous research conducted by Smith-Jackson et al. (2001) that are relevant to the proposed study is as follows:

User Requirements Specification:

1. Content

Users reported the general purposes for which they would refer to a healthcare information website as follows:

1. Where to find other links and information sources
2. Definitions of diseases
3. Definitions of medical terms
4. Drug uses, interactions, and side effects
5. Health care benefits programs
6. FAQ option
7. Herbal remedies
8. Information about health care plans and physicians
9. Prevention information
10. Directory system
11. Local events bulletin board (p. 3)

2. Specifications for Navigation, Selection Devices, Design, Layout and Other Features

General navigation tools

b.) Users expressed dissatisfaction with scrolling demands. Scrolling to view information should be minimized.
c.) Website features such as search engines, links, and menu items should be designed such that they are clearly visible and viewed without scrolling.
d.) Horizontal scrolling should be avoided.
e.) To assist with navigation, users should be given pointers to get to special functions on the Web site. For instance, if a Web site allows users to automatically email information to a friend, that function should be highlighted with a call-out or instruction box with brief text on how to perform that function.
f.) Selection or navigation devices such as drop down boxes or links shall be spaced apart to minimize confusion.
g.) Buttons used to navigate, submit, or activate must be easy to view. Redundant coding using size and color enhance usability.
h.) Attention-capture strategies such as large labels or high contrast colors should be used to direct users' attention to special functions or features. (p. 6)

Links

a.) Users preferred to have links to other Web sites. Links shall be easily distinguishable from other text and viewable from the initial page without scrolling.
b.) Novices relied heavily on links to conduct searches for various texts. Links, however, should be clearly distinguishable from the text. (p. 7)

Information Design and Terminology

b.) Information layout should avoid clutter or "business".
c.) Information shall be presented in a column format to avoid the need to scroll horizontally.
d.) In addition to avoiding clutter, links, menus, and function or navigation links embedded in a page shall be easily distinguishable from other text. (p. 7)

Fonts
b.) Font sizes shall be a minimum of 12 point (including links).
c.) Most users preferred not to have to enlarge fonts. Fonts should be clearly visible (> 10 point) without requiring enlargement by users.
d.) However, font sizes should not be excessively large. Users reported a preference for fonts that were "average", but not too small.
e.) ALL CAPS should be used for labels on Tabs. (p.8)

Color
a.) Colors shall be high contrast colors.
b.) High contrast shades of the same color shall not be used (e.g., dark green on light green).
c.) Overall use of color should be minimized. Users preferred a minimalist approach to color usage.
d.) The initial design involving tabs with white letters/font on a background that faded from blue to white was difficult to see. The background and font should be one consistent and high-contrast color combination. (p. 9)

Guidelines Derived from the Online Needs Assessment Tool:

1. Task Anxiety

The types of Web links selected should be those that are easy to use. This will reduce the anxiety associated with exploring new Web sites (p. 21).

2. Suggested Guidelines from Computer Usage Data

Simple navigation tools or tips should be provided to users when they enter the portal page (p. 22).

Online Needs Assessment General Specifications:

1. Use of Online Needs Assessment

a) The online tool must be tested in all browser environments to ensure usability.
b) Developers should be aware that many participants may complete the survey during non-business hours. Thus, some type of technical support during these hours may be needed.
c) Users provided solicited feedback during phase 2 implementation regarding the design of the online tool. Feedback included: problems with the length of time to complete the tool and problems understanding how to go back within a frame (p. 23).

2. Assessment Questions

a) Use open fields, check boxes, and radio buttons as needed
b) Accessibility issues related to vision should be addressed more thoroughly. Texts from the users' responses are provided below:

   Due to vision problems, the screen items may not be clearly visible. When I receive error messages, they are hard to understand.

   The text was difficult to read. Some of the vocabulary needed to be interpreted (browsers, input, search engines, etc.). The introduction to this survey mentioned the skill level of the computer user; I feel that someone with no experience or a beginner would have difficulty completing this survey. I used assistance in completing the survey due to vision problems.
c) Minimize the need to enter information in fields. Reduce typing demands.
d) Avoid technical jargon. Words such as "Browser" should be followed with an explanatory phrase or parenthetical definition.
e) Use automatic jumps to the next field when field entry is required (i.e, when field is full or maximum characters have been reached). This should be applied to short field only. "Essay" type fields should not move automatically.
f) Provide alternative key control messages so that users who have motor difficulty can select check boxes and radio bottoms without reliance on a mouse.
g) Questions related to information density should refer to "clutter" to assess user's opinions of information density.
h) Questions referring to "ease-of-use" or "simplicity" are easily understood by users and should be included as much as possible to assess design needs and preferences.
i) Questions referring to information quality or informativeness should be included in the assessment. In user descriptions of various Web sites, informativeness or information wealth were routinely mentioned.
j) Questions and response alternatives should be developed using information from focus groups.

k) “Other” or open fields should be used for questions with response alternatives that are not comprehensive (p. 24).

3. Embedded Questions and Survey Usability

a) Users were confused by embedded links in the interactive focus groups (C& D). Thus, frames should be used so respondents can view a Web site and answer a question without having to move to another page. Most of the online survey participants reported that frames were easy to use.
b) The online assessment tool instrument used a navigation button marked “Next” at the bottom of each page. “Next” and other navigation buttons should be used whenever possible.
c) One hundred percent of the respondents reported that it was “Very Easy” to select response circles (radio buttons). Due to the constraints of the software package used to create the online instrument, the radio buttons could not be enlarged. Respondents did not seem to have difficulty with the size or positioning the mouse.

APPENDIX D

THE PARTICIPANT CATEGORIZATION TOOL

Since there is only one published tool or method used in the area of user classification, which was a general description of three main dimensions on which users' experience differs (Nielsen, 1993a), a global classification scheme will be used. Experience level classification will be done by using a prescreening survey and a classification scheme. The Prescreening survey will be administered by email or hard copy. The survey and instructions for completion are provided on the next page.

Please answer the following questions from (1 to 14) by typing yes or no, in each of the answer boxes provided on the left side of each of the questions.

THE QUESTIONS	THE ANSWERS
1) Do you get connected to the World Wide Web?	
2) Do you use a mail reader (Ex: Microsoft Outlook, Outlook Express, Netscape, Eudora, AOL, Earthlink, Etc.)?	
3) Do you use a Web Browser (Ex: AOL, Netscape, Internet Explorer, Etc.)?	
4) Do you use Web search engines (Ex: Yahoo, Altavista, Etc.)?	
5) Do you perform any activities such as composing, sending, replying or forwarding email?	
6) Do you use the computer's spell check function while writing an email?	
7) Do you attach files or folders to the email body?	
8) Do you use any word processing software? (Ex: Microsoft Word, Wordpad, Word Perfect, Etc?	
9) Do you use Spread Sheets (Ex: Microsoft Excel, Etc)?	
10) Do you use presentation software (Ex: Microsoft PowerPoint, Hypercard, Etc)?	
11) Do you use any instant messengers (Ex: Aol IM, Yahoo IM, Msn Im, Etc)?	
12) Have you developed software or software applications?	
13) Have you developed Web sites?	
14) Have you ever completed an online survey?	
15) How frequently do you use the computer? (Please select a, b or c from the next column)	a) Less than 10 times a year or less than 3 times a month b) Less than 3-4 days a week c) Everyday

The classification criterion is based on the answers to the prescreening questionnaire, which will define the experience levels as described in Table 6.

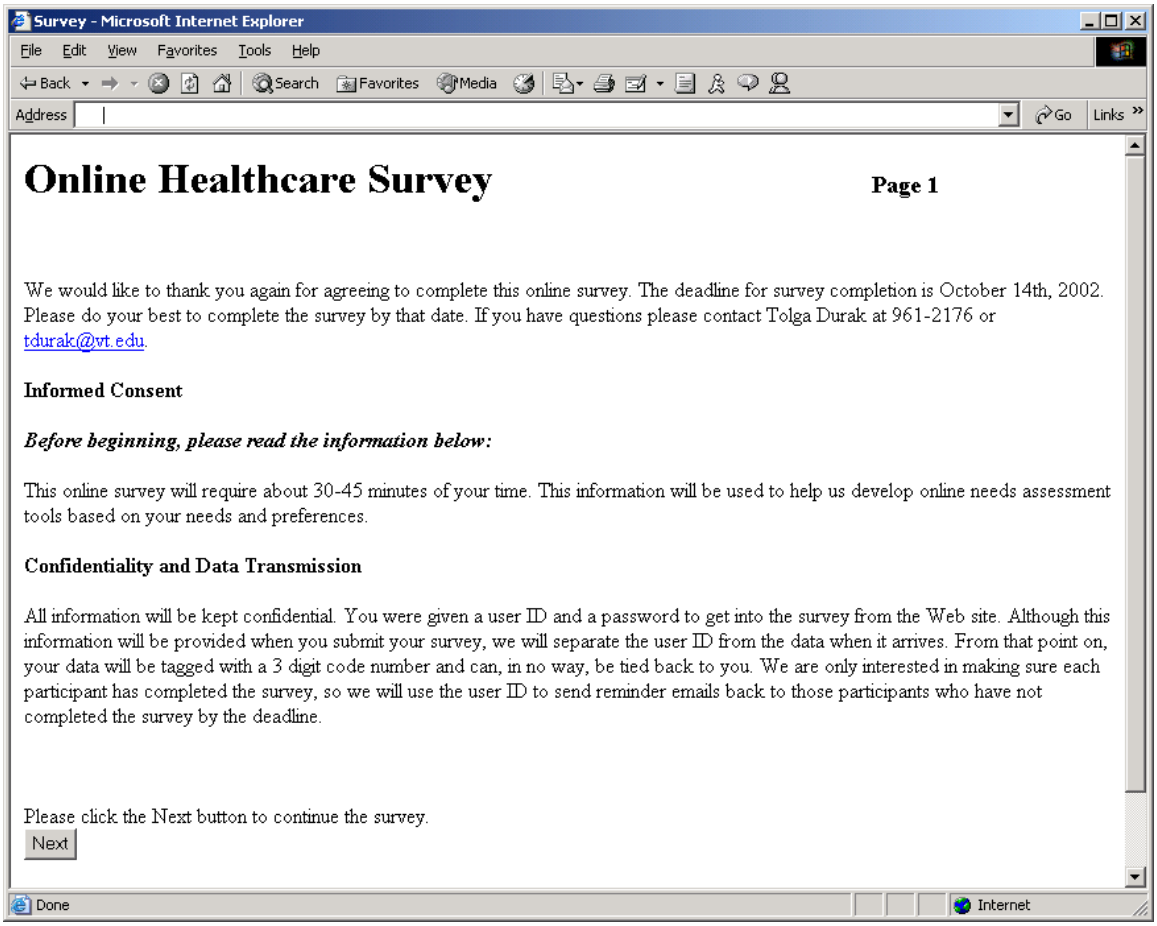
Table 6: The Table of Classification Criterion.

Novice:	If the total amount of yes answers are less than 5 (for questions from 1 to 14) and the answer to question 15 is either a or b
Intermediate:	If the total amount of yes answers are equal to 5 or less than 11 (for questions from 1 to 14) and the answer to question 15 is b or c.
Expert:	If the total amount of yes answers are equal to or more than 11 (for questions form 1 to 14) and the answer to question 15 is b or c.

APPENDIX E

PROTOTYPE B:

Screen shots from Prototype B, which was developed for Study 1, is provided below in a sequential order:



Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address Go Links

Online Healthcare Survey

Page 4

Part I: Online Questionnaire

The following is an Online Questionnaire composed of general questions on computers, Internet usage and preferences regarding healthcare websites. Please answer the following 11 questions of this questionnaire. If you have any difficulties, please do not hesitate to contact Tolga Durak at 961-2176 or tdedu@hotmail.com.

Please click the Next button to continue the survey.

Next

Done My Computer

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address Go Links

Online Healthcare Survey

Page 5

1. For which of the following do you use the Internet? (Check all that apply. Position the arrow over the square you want to check and click the mouse.)

- ☐ Check, send, forward, write, reply to an email
- ☐ Participating in online Chats
- ☐ Searching for information
- ☐ Downloading new software and updates from the Internet
- ☐ Checking stock quotes, news, weather, or other "daily" information
- ☐ Entertainment: Playing games and listening to music
- ☐ Other

2. If you selected "Other" in question 2 above, please type in a description below: (Position the cursor in the box and begin typing)

Please click the Next button to continue the survey.

Next

Done Internet

Survey - Microsoft Internet Explorer

Online Healthcare Survey

Page 10

8. How would you rate each of the following topics, which would appear on a healthcare website. (Position the arrow over the circle beside the response you would like to select. Click the mouse to select the answer that best matches your opinion)

(1 - Extremely Unimportant; 3 - Unimportant; 5 - Neutral; 7 - Important; 9 - Extremely Important)

Only 5 of the anchors are labeled, however you can assign any of the 9 numbers provided on the scale below.

	1	2	3	4	5	6	7	8	9
a) Research news, special reports, and clinical trials	○	○	○	○	○	○	○	○	○
b) Hospice services and bereavement counseling	○	○	○	○	○	○	○	○	○
c) Chats, aging healthy chats, bulletin board system (or discussion groups)	○	○	○	○	○	○	○	○	○
d) Self-Care services such as wellness tests, diet and fitness calculators	○	○	○	○	○	○	○	○	○
e) Opportunities to give feedback about the design for the Web site content	○	○	○	○	○	○	○	○	○

Please click the Next button to continue the survey.

Next

Survey - Microsoft Internet Explorer

Online Healthcare Survey

Page 13

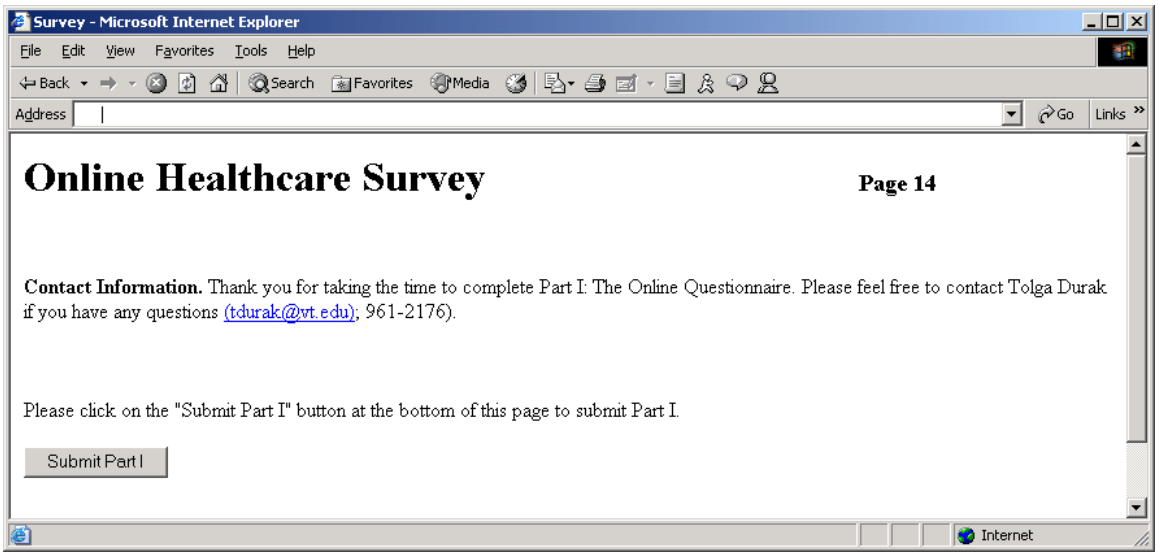
13. If a FAQ section were included on a Web site, which of the following would you prefer?

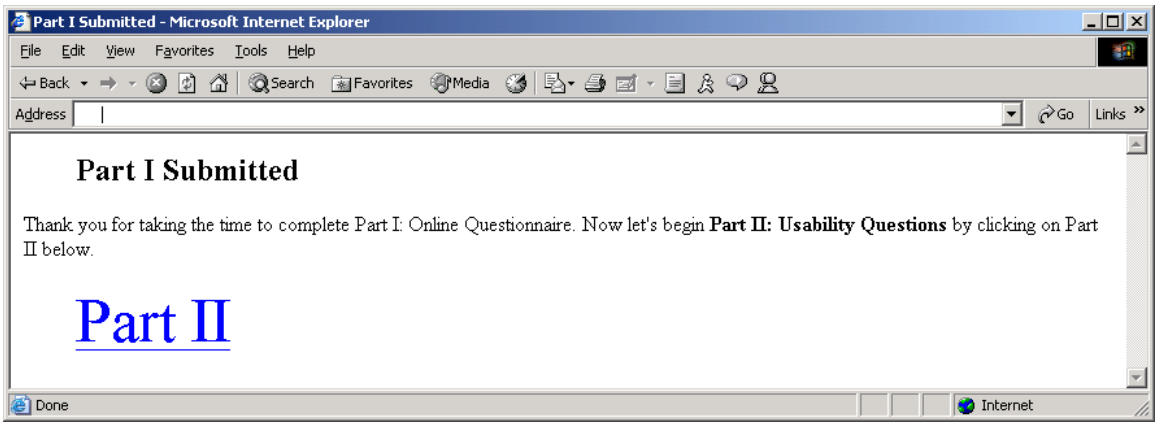
(Position the arrow over the circle beside the response you would like to select. Click the mouse to select the answer that best matches your opinion.)

- ○ A FAQ section that allows me to ask any question or any topic
- ○ A FAQ section that restricts questions to just a few topics (such as medications or disease symptoms).
- ○ A FAQ section that restricts questions to one topic only (such as medications or disease symptoms).

Please click the Next button to continue the survey.

Next


Survey - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address
Go Links
Online Healthcare Survey
Page 14
Contact Information. Thank you for taking the time to complete Part I: The Online Questionnaire. Please feel free to contact Tolga Durak if you have any questions (tdurak@vt.edu); 961-2176).
Please click on the "Submit Part I" button at the bottom of this page to submit Part I.
Submit Part I
Internet


Part I Submitted - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address
Go Links
Part I Submitted
Thank you for taking the time to complete Part I: Online Questionnaire. Now let's begin Part II: Usability Questions by clicking on Part II below.
Part II
Done
Internet

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://ace.ise.vt.edu/tdurak/part2_frameset.htm Go Links

Online Healthcare Survey

Page 15

Part II: Usability Questions

In this part, you are asked to answer a total of 20 usability questions followed by two open-ended question. These questions are specifically related to the Online Questionnaire, which you have completed in Part I. This questionnaire is also displayed in the lower part of this window as a link you can open. You are free to explore the Online Questionnaire to answer each question. When you want to close the Online Questionnaire window and return to Part II, click on the link named "Return to Part II". If you have any difficulties, please do not hesitate to contact Tolga Durak at 961-2176 or tdurak@vt.edu.

Identification Number.

Please type your 2-digit identification number into the box below. This is the number you received in a previous email. (Position the cursor in the box and begin typing.

Please click the Next button to continue the survey.

Next

Show Part I: Online Questionnaire This navigation bar is provided to make it possible to view Part I. In order to return to Part II, click on the "X" button on the top right hand corner of your browser screen.

Internet

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address Go Links

Online Healthcare Survey

Page 15

Part II: Usability Questions

In this part, you are asked to answer a total of 20 usability questions followed by two open-ended question. These questions are specifically related to the Online Questionnaire, which you have completed in Part I. This questionnaire is also displayed in the lower part of this window as a link you can open. You are free to explore the Online Questionnaire to answer each question. When you want to close the Online Questionnaire window and return to Part II, click on the link named "Return to Part II". If you have any difficulties, please do not hesitate to contact Tolga Durak at 961-2176 or tdurak@vt.edu.

Identification Number.

Please type your 2-digit identification number into the box below. This is the number you received in a previous email. (Position the cursor in the box and begin typing.

Please click the Next button to continue the survey.

Next

Show Part I: Online Questionnaire This navigation bar is provided to make it possible to view Part I. In order to return to Part II, click on the "X" button on the top right hand corner of your browser screen.

Done Internet

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address Go Links

Online Healthcare Survey

Page 17

2. Based on your interaction with the Online Questionnaire during Part I, please tell us how would you rate each of the following questions. (Position the arrow over the circle beside the response you would like to select. Click the mouse to select the answer that best matches your opinion)

(1 - Very Difficult; 3 - Difficult; 5 - Neutral; 7 - Somewhat Easy; 9 - Very Easy)

Only 5 of the anchors are labeled, however you can assign any of the 9 numbers provided on the scale below.

	1	2	3	4	5	6	7	8	9
1) Going to the next page	○	○	○	○	○	○	○	○	○
2) Reading the questions	○	○	○	○	○	○	○	○	○
3) Operating my input devices (mouse, trackball, stylus, keyboard) to select responses or insert responses	○	○	○	○	○	○	○	○	○
4) Understanding the concept of an "Online Questionnaire"	○	○	○	○	○	○	○	○	○

Please click the Next button to continue the survey.

Next

Show Part I: Online Questionnaire This navigation bar is provided to make it possible to view Part I. In order to return to Part II, click on the "X" button on the top right hand corner of your browser screen.

Done Internet

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address Go Links

Online Healthcare Survey

Page 22

Part III: Guideline Questions

In this part, you are asked to answer a total of 20 rating and two open-ended questions. These questions are specifically related to the usability guidelines used to develop the Online Questionnaire, which you have completed in Part I. This questionnaire is also displayed in the lower part of this window as a link you can open. You are free to explore the Online Questionnaire to answer each question. When you want to close the Online Questionnaire window and return to Part II, click on the link named "Return to Part II". For each of the 20 rating questions, you are asked to answer in two different aspects: Agreement and Importance.

On the Agreement scale, you can assign a number on a scale from 1 to 9, where 1 represents "Strongly Disagree", 3 represents "Disagree", 5 represents "Neutral", 7 represents "Agree", and 9 represents "Strongly Agree".

On the Importance scale, you can assign a number on a scale from 1 to 9, where 1 represents "Extremely Unimportant", 3 represents "Unimportant", 5 represents "Neutral", 7 represents "Important", and 9 represents "Extremely Important".

If you have any difficulties, please do not hesitate to contact Tolga Durak at 961-2176 or tdurak@vt.edu.

IdentificationNumber.

Please type your 2-digit identification number into the box below. This is the number you received in a previous email. (Position the cursor in the box and begin typing.

Please click the Next button to continue the survey.

Next

Show Part I: Online Questionnaire This navigation bar is provided to make it possible to view Part I. In order to return to Part II, click on the "X" button on the top right hand corner of your browser screen.

Done Internet

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address Go Links

Online Healthcare Survey

Page 23

In. this section all of the questions include more than one feature. When rating those questions, please rate the features as one group, rather than separately.

1. Color is properly used for the characters, foreground, background, and to highlight important information.

(Position the arrow over the circle beside the response you would like to select. Click the mouse to select the answer that best matches your opinion)

(**Agreement Scale** ==> 1 - Strongly Disagree; 3 - Disagree; 5 - Neutral; 7 - Agree; 9 - Strongly Agree)

(**Importance Scale** ==> 1 - Extremely Unimportant; 3 - Unimportant; 5 - Neutral; 7 - Important; 9 - Extremely Important)

Only 5 of the anchors are labeled, however you can assign any of the 9 numbers provided on the scale below.

	1	2	3	4	5	6	7	8	9
Agreement	○	○	○	○	○	○	○	○	○
Importance	○	○	○	○	○	○	○	○	○

Please click the Next button to continue the survey.

Next

Show Part I: Online Questionnaire This navigation bar is provided to make it possible to view Part I. In order to return to Part II, click on the "X" button on the top right hand corner of your browser screen.

Done Internet

APPENDIX F

A SAMPLE OF NEEDS ELICITATION QUESTIONS

A segment of Prototype B, which was developed for Study 1, is provided below:

Online Healthcare Survey

Page 7

5. Please examine the topics below. In your opinion, which of the following topics are important to include on a healthcare Web site (Check all that apply. Position the arrow over the square you want to check and click the mouse.) ?

- ☐ Where to find other links and information sources such as books and organizations that provide additional information
- ☐ Definitions of diseases
- ☐ Definitions of medical terms
- ☐ Drug uses, interactions (with food and herbal remedies), and side effects
- ☐ Health care benefits programs
- ☐ Frequently asked questions
- ☐ Herbal remedies
- ☐ Information about health care plans and associated physicians
- ☐ Prevention information such as nutrition and exercise information
- ☐ Directory system with hospital numbers, local ambulance numbers, maps to local hospitals and clinics
- ☐ Directory/bulletin board of local events, educational programs/workshops, health-related meetings

p (7)

APPENDIX G

A SAMPLE OF USABILITY QUESTIONS

On the next page a segment of an online needs assessment tool in the form of a Web-based questionnaire, which was administered in prior research by Smith-Jackson et al. (2001), is provided.

Survey: Part II

Sample Web Site Evaluation. This question refers to the above Web site. Before completing the next evaluation, please do the following:

1. Move the right vertical scroll bar inside the above frame and view the contents of the Web page.
2. Select an item (any item) from the circular menu at the top of the page.
3. Go back to the main page by clicking on the word **Home** or the **CareScout** picture, both of which are in the upper left corner of the **CareScout** web page.
4. Find the "email this page to a friend" link. It is in the lower right portion of the main page. Please email this page to anyone you wish.

You can use the back arrow at the top of your browser to move within this Web site. Also, two scroll bars will be on your screen. One is for the Carescout Web site and one is for this survey.

Identification number. Please type your 4-digit identification number into the box below. This is the number you received in a previous email. (Position the cursor in the box and begin typing.)

Question. Please rate how much you agree with the following statements about the Web site (www.CareScout.com) you just viewed.

	Agree Strongly	Agree Somewhat	Disagree Somewhat	Disagree Strongly
The "email to a friend" feature was easy to use.				
It was easy for me to move or navigate within the Web site.				
I liked the number of links provided on this Web site.				
I liked the design of the menu (round menu at the top of the page).				

(p. 1-5)

APPENDIX H

SAMPLE OF THE GUIDELINE EVALUATION QUESTIONS

On the next page a segment of an online needs assessment tool in the form of a Web-based questionnaire, which was administered in prior research by Williges, Smith-Jackson and Kwahk. (2001), is provided.

Part II: Online Questionnaire

You are asked to answer questions 1 through 21 on this page. Please scroll down to see the remainder of this page.

1. Color is properly used for the foreground, background, and to highlight important information.
Please select the number that best represents your opinion.

1A. Agreement

(1 - Strongly Disagree, 3 - Disagree, 5 - Neutral, 7 - Agree, 9 - Strongly Agree)

	1	2	3	4	5	6	7	8	9
Agreement	☐	☐	☐	☐	☐	☐	☐	☐	☐

1B. Importance

(1 - Extremely Unimportant, 3 - Unimportant, 5 - Neutral, 7 - Important, 9 - Extremely Important)

	1	2	3	4	5	6	7	8	9
Importance	☐	☐	☐	☐	☐	☐	☐	☐	☐

p (19-20)

APPENDIX I

GUIDELINES EXTRACTED FROM EXISTING RESEARCH RESULTS

The summary of the design guidelines extracted from existing research results and their origin are as follows:

Color, Background, Foreground

1) Maximize contrast between characters and background (Czaja, 1997).
2) Use easily discriminated color combinations (Czaja, 1997).
3) Avoid hard coding on font, size and color (Agelight, 2001).
4) Avoid distracting background elements (Agelight, 2001).
5) Avoid animation (Agelight, 2001).
6) Avoid exceptionally bright, fluorescence or vibrant colors (Agelight, 2001).
7) Avoid blue/yellow or red/green combinations (Agelight, 2001).
8) Use dark type on light background (Agelight, 2001).
9) Maximize contrast between foreground and background (Cast-Bobby, 2001).
10) Design in black and white; add color for emphasis (Hyatt, 2000).
11) Don't rely on color alone for information cue (Hyatt, 2000).
12) Maximize color contrast for people with color deficiencies (Hyatt, 2000).
13) Label each page with the site and page name (SPRY Foundation, 1999).
14) Use the same style throughout the site (SPRY Foundation, 1999).
15) Background color should be void of the blue/green range (Hawthorn, 200).
16) Graphics should be kept to a minimum or voided altogether (Hawthorn, 2000).
17) Incomplete, embedded or partial figures, charts, tables or reports in visual design applications should not be used (Humphrey, Kramer & Gore, 1994).
18) Use dark type on a light background, emphasizing high contrast (Ellis, Kurniawan, 2000, p 273-274).

Font

1) Use legible, commonly-used font throughout the site (Agelight, 2001).
2) Use 12 to 14-point font sizes, and two points larger for titles (Agelight, 2001).
3) Avoid mixed and decorative fonts (Agelight, 2001).
4) Use bold type weight to emphasize (Agelight, 2001).
5) Avoid all capital letters even in headings (Agelight, 2001).
6) Use sans serif fonts, and avoid stylized fonts. (SPRY Foundation, 1999).
7) Use medium level type weight (SPRY Foundation, 1999).
8) Text size should be set in the 12 – 14 point size range (Hawthorn, 2000).
9) Use a sans serif font, such as Arial or Helvetica (Ellis, Kurniawan, 2000, p 273-274).

General Navigation Tools

1) Avoid small targets and characters (Czaja, 1997).
2) Clearly label keys (Czaja, 1997).
3) Maximize size of icons (Czaja, 1997).
4) Use meaningful icons and label them (Czaja, 1997).
5) Area cursor and sticky icon interaction techniques should be used at the same time (Worden, Walker, Bharat, Hudson, 1997).
6) Scrolling to view information should be minimized. Horizontal scrolling should be avoided (Smith-Jackson et al., 2001, p.6).
7) To assist with navigation, users should be given pointers to get to special functions on the Web site. For instance, if a Web site allows users to automatically email information to a friend, that function should be highlighted with a call-out or instruction box with brief text on how to perform that function (Smith-Jackson et al., 2001, p.6).
8) Selection or navigation devices such as drop down boxes or links shall be spaced apart to minimize confusion (Smith-Jackson et al., 2001, p.6).
9) Buttons used to navigate, submit, or activate must be easy to view. Redundant coding using size and color enhance usability (Smith-Jackson et al., 2001, p.6).

Information Design and Terminology

1) Minimize irrelevant information (Czaja, 1997).
2) Present information in consistent locations (Czaja, 1997).
3) Adhere to principles of perceptual organizations (Czaja, 1997).
4) Highlight important screen information (Czaja, 1997).
5) Avoid complex command languages (Czaja, 1997).
6) Use consistent operating procedures (Czaja, 1997).
7) Balance the use of open space (Agelight, 2001).
8) Leave right margin for various window/display sizes (Agelight, 2001).
9) Break topics down into succinct pages (Agelight, 2001).
10) Avoid reduced kerning or condensed spacing (Agelight, 2001).
11) Clarify natural language usage (Cast-Bobby, 2001).
12) Ensure that documents are clear and simple (Cast-Bobby, 2001).
13) Use the same style throughout the site (SPRY Foundation, 1999).
14) Leave enough space between paragraphs (Health Canada, 1999).
15) The visual field should be placed in narrow view (Hawthorn, 2000).
16) Distorted spacing should be avoided; the block justification style should not be used (Shaw, 1990).
17) Design your application consistent with software applications commonly used (Cooper, Lee, Goska, Anderson, Gay, Fickes & Fisk 1992).
18) Plan screen sizes to fit in lower resolution display setting, at least 800*600 and perhaps even 640*480 for monitors up to a 17-inch screen size. Plan for users to browse with a single maximized (full-screen) window, and minimize the use of links and references to new browser windows. (Ellis, Kurniawan, 2000, p 273-274).

19) Information layout should avoid clutter or “business” (Smith-Jackson et al., 2001, p.7).
20) Information shall be presented in a column format to avoid the need to scroll horizontally (Smith-Jackson et al., 2001, p.7).
21) In addition to avoiding clutter, links, menus, and function or navigation links embedded in a page shall be easily distinguishable from other text (Smith-Jackson et al., 2001, p.7).

Links

1) Avoid link names like “click here” (Apple Computers, 2001).
2) Avoid links requiring precise mouse control (SPRY Foundation, 1999).
3) Make sure that links (a) are placed where they are easy to see, (b) are fairly large (the size settled on by the PD team here was 180 pixels * 22 pixels for a graphic button), and (c) have plenty of dead space around them to prevent accidental selection (Ellis, Kurniawan, 2000, p 273-274).
4) Users preferred to have links to other Web sites. Links shall be easily distinguishable from other text and viewable from the initial page without scrolling (Smith-Jackson et al., 2001, p.7).
5) Novices relied heavily on links to conduct searches for various texts. Links, however, should be clearly distinguishable from the text (Smith-Jackson et al., 2001, p.7).

Help & Online/Offline-Support

1) Provide easy to use help (Czaja, 1997).
2) Offer different ways to fill out forms (Apple Computers, 2001)
3) Provide alternative mechanisms for on-line forms (City of San Jose, 1998)

APPENDIX J

SUPPLEMENTAL RESULTS FOR STUDY 1:

a) Close-Ended Questions with check-box alternatives:

In Part I, a total of five close-ended questions (Questions 1,5,11,12, and 13) were used. For each of these questions and their response alternatives, a separate 2*3 Chi-Square Test was performed. Chi-Square Tests were performed for experience level categories as well as gender, due to the equal number of male and female participants. In addition, for several response alternatives, Chi-Square test was identified as an invalid test due to having fewer than 5 elements per a response cell. For these response alternatives, separate Mantel-Haenszel Chi-Square tests were performed in order to determine the significant differences.

The first question, which was identified as Question 1 in Part I, was used to elicit the frequency for each Internet use motive. For this question, all of the participants chose one or more of the response alternatives (as shown in Table 2) for their Internet use motive(s). The results provided frequency for each motive (as shown in Table 2). However, gender and experience level categories were found to have no significant effect ($p > .05$) on Internet use motives represented with the response alternatives.

Table 2: The Internet Use

Response Alternatives (R. A.)	I USE the internet for this item	I DO NOT USE the internet for this item
Check, send, forward, write, reply to an email	100%	0%
Participating in online Chats	53%	47%
Searching for information	100%	0%
Downloading new software and updates from the Internet	93%	7%
Checking stock quotes, news, weather, or other "daily" information	76%	24%
Entertainment: Playing games and listening to music	83%	17%

The second question was identified as Question 5 in Part I. This question asked the participants to select the content items that are important and should be included in a healthcare Web site. All thirty of the participants chose one or more of the response alternatives (as shown in Table 3) as important content item(s). The results provided the frequency for each content item (as shown in Table 3). However, gender and experience level categories were found to have no significant effect ($p > .05$) on important content items represented with response alternatives.

Table 3: Important Content Items

Response Alternatives	IMPORTANT to include	NOT IMPORTANT to include
Where to find other links and information sources such as books and organizations that provide additional information	100%	0%
Definitions of diseases	90%	10%
Definitions of medical terms	96%	4%
Drug uses, interactions (with food and herbal remedies), and side effects	93%	7%
Health care benefits programs	100%	0%
Frequently asked questions	96%	4%
Herbal remedies	73%	17%
Information about health care plans and associated physicians	100%	0%
Prevention information such as nutrition and exercise information	93%	7%
Directory system with hospital numbers, local ambulance numbers, maps to local hospitals and clinics	90%	10%
Directory/bulletin board of local events, educational programs/workshops, health-related meetings	66%	34%

The third and the forth questions, which were identified as Questions 11 and 12 in Part I, was used to elicit information on FAQ (Frequently Asked Question) section usefulness. The results showed that 80% of the participants (24 out of 30 participants) have used and the remaining 20% of the participants have not used the FAQ section. In addition, 63% of the participants (19 out of 30 participants) thought FAQ section would

be useful to them, 6% of the participants did not think that the FAQ section would be useful to them and the remaining 31% did not have any opinion. However, gender and experience level categories were found to have no significant effect on ($p > .05$) on FAQ section usefulness.

The last question, which was identified as Question 13 of Part I, was used to elicit information on FAQ format preference. This question asked the participants to select from one of the three response alternatives and indicate their preferred FAQ format. All thirty of the participants selected one FAQ format as their preference (as shown in Table 4). The results provided the frequency for each FAQ format (as shown in Table 4). However, gender and experience level categories were found to have no significant effect ($p > .05$) on the FAQ formats represented with response alternatives.

Table 4: FAQ Format

Response Alternative	I WOULD prefer to have	I WOULD NOT prefer to have
A FAQ Section that allows me to ask any question or any topic	76%	24%
A FAQ Section that restricts questions to just a few topics.	24%	76%
A FAQ Section that restricts questions to one topic only.	0%	100%

Subsequent to data analysis and interpretation, guideline extraction methodology was applied to the results of these five questions. However, neither one of these questions or their response alternatives met the guideline extraction criteria (see Section 3.1.4 for more details on the criteria). Therefore, no item was excluded from the response alternatives and no guideline was developed.

<u>b) Close-Ended Questions with radio-button alternatives (9-point Likert-Type scale questions):</u>

The 9-point Likert-Type scale questions were used to have participants rate the importance of various healthcare contents and features that were presented with response alternatives. The mean ratings for each item are provided in Table 5.

Table 5: Healthcare Contents and Features

Item	Mean Ratings	SD
Directory system with hospital numbers, etc., email addresses of physicians, phone numbers of physicians, maps	**7.17**	**0.23**
Educational programs, events, health-related meetings in local area	**6.83**	**0.26**
Exercise/nutrition/prevention information	**6.73**	**0.26**
Pharmaceutical information: Drug information, interactions, side effects	**6.73**	**0.32**
Emergency actions or special actions in case of death occurring in home	**6.63**	**0.35**
Health care plans and physicians associated with each health plan (regional, local).	**6.53**	**0.26**
Symptoms and treatments for diseases	**6.53**	**0.33**
Information on which physicians in local area accept Medicare	**6.43**	**0.51**
Search engine with easy filter system	**6.37**	**0.98**
Health care benefit programs: Caregivers, assisted living, hospices, durable medical equipment, home health care	**6.33**	**1.12**
Definitions of diseases and medical terms	**6.10**	**1.12**
Sites to purchase medications	**6.07**	**1.12**
Knowing the sponsor(s) of a Web site you are using (for instance, insurance company sponsors, pharmaceutical sponsors)?	**5.63**	**1.32**
Where to find other information sources such as other links, books, organizations	**5.63**	**1.42**
Frequently asked questions option	**5.50**	**1.12**
Self-Care services such as wellness tests, diet and fitness calculators	**5.30**	**1.32**
Opportunities to give feedback about the design for the Web site content	**5.20**	**1.21**
Hospice services and bereavement counseling	**5.00**	**1.21**
Research news, special reports, and clinical trials	**4.77**	**1.53**
Herbal remedies	**4.57**	**1.42**
Chats, aging healthy chats, bulletin board system (or discussion groups)	**4.10**	**1.42**

Subsequent to calculating the mean ratings, guideline extraction methodology was applied to the results of these four questions. However, neither one of the response alternatives met the guideline extraction criteria (see Section 3.1.4 for more details on the criteria). Therefore, no guideline was developed.

In addition to the guideline extraction, 2*3 ANOVA was performed for each of the response alternatives and for completion times in order to determine the significant differences. Each ANOVA test was performed for experience level categories as well as gender, due to the equal number of male and female participants (see Appendix R for ANOVA results). The results of these analyses indicated two significantly different results. In addition to completion time, only one response alternative showed significantly different results.

For completion time, the following significant effect was found:

- There was a significant difference between the mean completion time for female and male participant groups ($F\ (1,24) = 70.08,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean completion time for female participants ($M = 9.00,\ SD = 1.36$) was significantly higher than the mean completion time for male participants ($M = 5.13,\ SD = 1.19$).

For the content item "Educational programs, events, health-related meetings in the local area" (response alternative 5 of question 7), the following significant effect was found:

- There was a significant difference between the mean rating value for female and male participant groups ($F\ (1,24) = 10.47,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean rating value for male participants ($M = 7.93,\ SD = 1.39$) was significantly higher than the mean rating value for female participants ($M = 5.73,\ SD = 2.22$). Male participants expressed a significantly stronger preference towards including educational programs, events, health-related meetings in the local area in a healthcare website.

The Cronbach coefficient alpha value was *.94* and the range of item-to-total correlation was *.29* – *.82*. The correlation coefficients among the 21 response alternatives were all positive and statistically significant at the *.05* level. The correlation coefficients ranged from a low *.29* to a high *.82*, indicating the degree of the relationships between the response alternatives. The lowest reliable variance of the total scores was at least 93.6%. In their study, Gable and Wolf (1993) suggest that the alpha reliability should at least be a .70 and a value higher than .80 would be very satisfactory to achieve. According to this finding, each response alternative had adequate internal consistency and there was no need to further improve the reliability of these questions and their response alternatives.

c) Open-ended Questions, comments and suggestions sections:

In Part I, a total of four open-ended questions (Questions 2,3,4, and10) were used. The first three open-ended questions (Questions 2,3, and 4) were directed to elicit information on types of problems participants experienced when using computers and the W.W.W. The last open-ended question (Question 10) was directed to elicit any additional comments participants had regarding healthcare web site content. Not only all of these four questions provided supplemental demographic information, but also the responses to these questions had pertinent content for guideline extraction and ONAT iteration. For each of these questions, a separate content analysis was performed. Upon the completion of content analysis, for each question, final response categories and their response frequencies were identified and guidelines extraction methodology was applied.

The first question, which was identified as Question 2, was a follow-up question that was used to explore additional ways the participants used the Internet for. For this question, only one final response category was identified. Even though 15 participants expressed that they used the Internet for “Downloading/Checking class notes, and other class material”, the item was not relevant to the content of the Online Tool. Thus, the item was not developed into a response alternative and was not included in the list of the Internet Use Motives.

The second and third questions, which were identified as Question 3 and 4, were used to elicit the types of computer and Internet problems participants experienced. For computer problems, ten final response categories and for Internet problems, nine final response categories were identified (as shown in Table 7 and Table 8 respectively). Several of these final response categories supported existing guidelines that expressed the importance of simplicity, clarity, ease of use, information reliability and consistency. But none of the categories included new content or application area that could be developed into a new design guideline.

Table 7: Summary of Computer Problems

Problem	Frequency
Not being able to set or correct Internet settings	1
Not being able to understand certain procedures or commands	1
Slow Internet connection	1
Slow Processor	6
Computer Freezing/Locking up (Not enough memory)	9
Problems due to an old mouse that cannot function well	1
Unable to get rid of viruses	1
Computer starting to be behind time	1
Computer not turning on and off properly	1
Not being able to know how to use certain programs	2

Table 8: Summary of Internet Problems

Problem	Frequency
Not being able to set or correct Internet settings	1
Not being able to find desired information while conducting searches	7
Slow Internet connection or not being able to maintain fast Internet connection	6
Websites not working or not being able to reach a site or its pages (coding or maintenance problems)	8
Slow opening of web pages	4
Pop-up advertisements freezing the computer	1
Words on a page not showing (coding or maintenance problems)	1
Not being able to find the sources of some online information or not being sure whether or not a certain information is reliable	1
Programs (such as emails or browsers) being terminated due to an error	1

The last question, which was identified as Question 10, was used to elicit any additional content participants would like to have on a healthcare information web site. For this question, the following nine final response categories were identified:

1) Comparisons to other healthcare programs
2) Catalogues that have a live online person you can ask questions to all the time
3) Information pertaining alternative treatments
4) List of certain symptoms of various diseases and a list of certain symptoms when person should go to see their doctor
5) Information on content validity
6) Price comparisons between different health insurance plans
7) Info about AIDS\HIV, and other STD's
8) Medical trials
9) Information on dietary supplements

Two of these categories included new content items that were not included in the current ONAT. However, all of these response categories received a frequency of one. Therefore, neither one of these additional content categories exceeded the overall score 0.4 (see Section 3.1.4 for more details on calculating the score). Therefore, no new content item was included in the next design iteration.

In addition, as her response to this question, one participant expressed her concern about to not being able to understand the concept of a “Healthcare Website”.

The overall score calculated for this concern did not exceed 0.4. Despite failing to meet the guidelines extraction criteria, due to several guidelines (that were used to develop Prototype B) that stated the importance of language clarity, comprehension and avoiding the use of technical terms without explanations, a definition of “Healthcare Website” was included. The definition was added to the beginning of Part I Introduction page.

APPENDIX K

THE SUMMARY OF HEURISTIC EVALUATION STEPS FOR STUDY 1

During the preliminary development activities, Prototype A was developed and improved to form Prototype B. During these activities, according to the design guideline set (see Appendix J for more details), several changes were made to Prototype A. The following are some of the main design changes presented in sequential order:

1) All the scales were changed to 9-point Likert-Type Scales.
2) The scales for part 1 were reversed so it's consistent throughout the tool
3) A requirement on the flyer as the participants must have visited 2-3 websites prior to the study was included so providing examples of healthcare websites was not necessary.
4) Before each question, a reminder sentence indicating that participants can assign any of the 9 anchors was placed.
5) The survey was divided into three parts in order to avoid having useless data from all of the 3 parts instead of only the single problematic section.
6) A sentence before part 3 indicating that participants should group features in the specific guidelines and rate accordingly was included.
7) In order to derive more information one open-ended question at the end of part 2 and part 3 were added.
8) Some design guidelines were not implemented due to technical limitations. Implementation of these guidelines required changing Online Tool coding with dreamweaver or flash… and this change would not allow the tool to work in some of the older browser versions.
9) Explanation phrases for the "next" and the "submit" buttons were not included due to the lack of space on each page. In addition, the results of the previous Online Needs Assessment Tool prototype evaluations showed that the older users didn't experience any difficulties without this phrase.
10) The help page link was excluded from the prototype. The Help icon was taking was taking up a considerable amount of space on each page, had compatibility problems with older browsers. Therefore, instead of having the help page as a link, Help tips, which included everything covered in the help page, were provided orally and as well as a document. The following page is the Help tips document provided to every participant.

Help Tips Document:

USER IDENTIFICATION NUMBER

Tolga Durak (the principle investigator) will provide you with this 2-digit number prior to the study via email.

If you have any problems or doubts regarding your user ID number, please do not continue the survey and contact Tolga Durak immediately (contact information is provided below).

NAVIGATION BUTTONS AND LINKS

If one of the links doesn't work, try to copy and paste the specific link.

The **"next"** button will enable you to move to the next page. You can always go back using the **"back"** button on your tool bar after using the **"next"** button.

The **"submit"** button will enable you to submit a specific survey or its' part. So you will not be able to go back to the previous page after clicking the **"submit"** button.

So please do not use the "submit" button unless you are completely satisfied with your responses. In case of any doubt or hesitation please do not use the "submit" button and contact Tolga Durak.

Please do not continue the survey and contact Tolga Durak immediately if you are experiencing any problems, or if something looks odd or sounds confusing.

CONTACT INFORMATION

If you have questions at any time about the project or the procedures, you may contact the principal investigator, Tolga Durak at 961-2176, tdedu@hotmail.com

After developing Prototype B, Pilot Testing was conducted. During the pilot testing 2 participants were asked to complete the demographic information questionnaire and Prototype B and answer the following questions after completing each section:

1) Did all the questionnaires work with your browser?
2) Did every page have a sufficient, understandable, readable and logical layout?
3) Were instructions such as hit “submit” button only once or use back button on the tool bar to go back to the previous page necessary?
4) Was the help provided (including reminders provided and the contact information) sufficient?
5) Were the part explanations, instructions on how to complete the survey, how to use the buttons, how to interact with the pop-up window sufficient?
6) Were the headings intuitive or should they be revised? Were they sufficient enough to provide track information?
7) Is the spacing in between text/between text and buttons sufficient?
8) Is the frame page clear and understandable?

The following results were gathered from the pilot testing:

- Every page was found to have a sufficient, understandable, readable and logical layout by both of the participants.
- Instructions such as hit “submit” button only once or use back button on the tool bar to go back to the previous page were not found to be necessary by both of the participants.
- The help provided (including reminders provided and the contact information) were found to be sufficient by both of the participants.
- The part explanations, instructions on how to complete the survey, how to use the buttons, how to interact with the pop-up window were found to be sufficient by both of the participants.
- The headings were found to be intuitive and sufficient enough to provide track information by both of the participants.
- The spacing in between text/between text and buttons were found to be sufficient by both of the participants.
- The frame page was found to be clear and understandable by both of the participants. However, both of the participants suggested making the information in the footer more concise in order to shrink the footer size.

In addition, during the pilot testing participants notified a few typos, punctuation, sentence alignment, anchor alignment, sentence justification and header justification mistakes due to the mistakes and mishaps in the source code of the prototype pages. Participants also expressed that they thought the frame page was taking up much screen space. The information in the footer should be more concise in order to shrink the footer size.

Upon the completion of pilot testing, the following changes were made to Prototype B:

- All source code mistakes and mishaps were eliminated and all the necessary cosmetic changes were made to the prototype pages.
- The frame page content and the footer size were decreased.
- Even though the questionnaire worked successfully with the browser choice of the pilot testing participants, the tools were not compatible with the older browsers. Because of the Java Script code integrated for the time stamp, problems occurred when older browsers (earlier versions of Internet Explorer, Netscape and AOL) were used. In order to avoid this incompatibility problem, the integrated Java Script code was substituted from the source code and participants were asked to record the time for each of the three sections.

APPENDIX L

INFORMED CONSENT FORM FOR YOUNGER PARTICIPANTS

Title of Project: Investigating the Usefulness of Age-Independent Online Needs Assessment Tools

Principal Investigators: Tolga Durak and Tonya L. Smith-Jackson, Ph.D.

PURPOSE OF PROJECT

You are invited to participate in a research project to design an online needs assessment tool to support health care interests of younger and older adults. You will be asked to participate one on-line survey session.

INFORMATION

The online survey will require one 20-30 minute session. These surveys can be completed either at the Assessment Cognitive Ergonomics Lab (Whittemore Hall room 519-A) at Virginia Polytechnic Institute and State University campus or at a location of your choice.

RISKS

There are no physical or emotional risks associated with this research project.

BENEFITS

At the end of this session, you will be provided with a copy of this form. At the bottom of this form, you will find contact information that can be used to contact the principal investigators after the research has been completed in order to receive information about the results.

CONFIDENTIALITY

The information gained in this research project will be kept strictly confidential. At no time will the researchers release the results of the study to anyone other than individuals working on the project without your written consent.

You will be identified only by a 3-digit study code. Data will be stored securely and will be made available only in the context of research publications and discussion. No reference will be made in oral or written reports that could link you to the data nor will you ever be identified as a participant in the project.

COMPENSATION

Your participation will benefit older and younger adults who need access to quality health care information. Personal satisfaction can result from knowing you have contributed to a worthwhile effort. At the end of this study, your effort will be compensated with your preference of 1 research credit or $7.50.

FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time without penalty.

APPROVAL

This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects At Virginia Polytechnic Institute and State University and by the Department of Industrial and Systems Engineering.

PARTICIPANT'S RESPONSIBILITIES

It is very important that you keep the activities and information discussed confidential, since others will be participating in this research.

QUESTIONS

If you have questions, or do not understand information on this form, please feel free to ask them now.

PARTICIPANT'S PERMISSION

I have read and understand the Informed Consent and conditions of this project. I have had all questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project.

If I participate, I may withdraw at any time without penalty.

Signature

_____

Date _____

CONTACT

If you have questions at any time about the project or the procedures, you may contact the principal investigators, Tolga Durak at 552-4777, tdurak@vt.edu or Dr. Tonya Smith-Jackson at 231-4119, smithjack@vt.edu.

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant have been violated during the course of this project, you may contact Dr. D. M. Moore, Chair of the Institutional Review Board Research Division at 231-4991 or moored@vt.edu.

APPENDIX M

INFORMED CONSENT FORM FOR OLDER PARTICIPANTS

Title of Project: Investigating the Usefulness of Age-Independent Online Needs Assessment Tools

Principal Investigators: Tolga Durak and Tonya L. Smith-Jackson, Ph.D.

PURPOSE OF PROJECT

You are invited to participate in a research project to design an online needs assessment tool to support health care interests of younger and older adults. You will be asked to participate in one of the on-line surveys.

INFORMATION

The online survey will require one 30-45 minute session. These surveys can be completed either at the senior center at the Blacksburg Parks and Recreation Center, at the Assessment Cognitive Ergonomics Lab (Whittemore Hall room 519-A) at Virginia Polytechnic Institute and State University campus, or any other convenient location.

RISKS

There are no physical or emotional risks associated with this research project.

BENEFITS

At the end of this session, you will be provided with a copy of this form. At the bottom of this form, you will find contact information that can be used to contact the principal investigators after the research has been completed in order to receive information about the results.

CONFIDENTIALITY

The information gained in this research project will be kept strictly confidential. At no time will the researchers release the results of the study to anyone other than individuals working on the project without your written consent.

You will be identified only by a 3-digit study code. Data will be stored securely and will be made available only in the context of research publications and discussion. No reference will be made in oral or written reports that could link you to the data nor will you ever be identified as a participant in the project.

COMPENSATION

Your participation will benefit older and younger adults who need access to quality health care information. Personal satisfaction can result from knowing you have contributed to a worthwhile effort. At the end of this study, your effort will be compensated with your choice of Virginia Tech paraphernalia.

FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time without penalty.

APPROVAL

This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects At Virginia Polytechnic Institute and State University and by the Department of Industrial and Systems Engineering.

PARTICIPANT'S RESPONSIBILITIES

It is very important that you keep the activities and information discussed confidential, since others will be participating in this research.

QUESTIONS

If you have questions, or do not understand information on this form, please feel free to ask them now.

PARTICIPANT'S PERMISSION

I have read and understand the Informed Consent and conditions of this project. I have had all questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project.

If I participate, I may withdraw at any time without penalty.

Signature

_____

Date _____

CONTACT

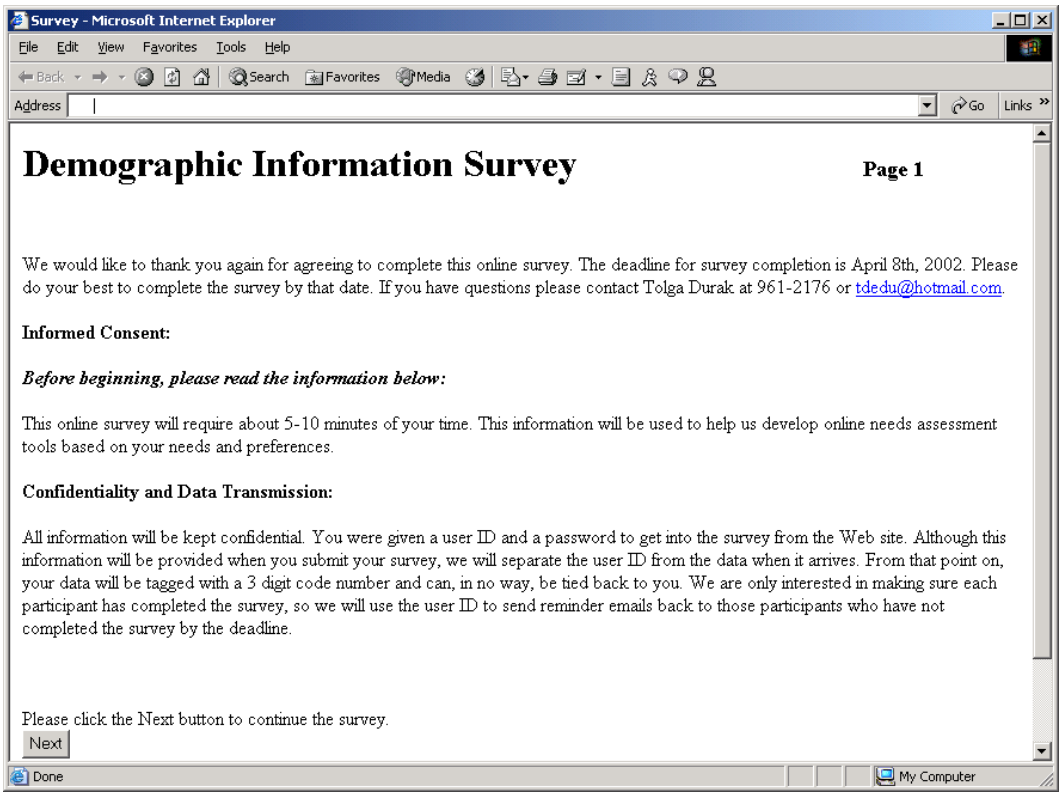
If you have questions at any time about the project or the procedures, you may contact the principal investigators, Tolga Durak at 552-4777, tdurak@vt.edu or Dr. Tonya Smith-Jackson at 231-4119, smithjack@vt.edu.

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant have been violated during the course of this project, you may contact Dr. D. M. Moore, Chair of the Institutional Review Board Research Division at 231-4991 or moored@vt.edu.

APPENDIX N

DEMOGRAPHIC INFORMATION QUESTIONNAIRE:

Screen shots from demographic Information Questionnaire, which was developed for Study 1 and 2, is provided below in a sequential order:



Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address Go Links

Demographic Information Survey **Page 1**

We would like to thank you again for agreeing to complete this online survey. The deadline for survey completion is April 8th, 2002. Please do your best to complete the survey by that date. If you have questions please contact Tolga Durak at 961-2176 or tdedu@hotmail.com.

Informed Consent:

Before beginning, please read the information below:

This online survey will require about 5-10 minutes of your time. This information will be used to help us develop online needs assessment tools based on your needs and preferences.

Confidentiality and Data Transmission:

All information will be kept confidential. You were given a user ID and a password to get into the survey from the Web site. Although this information will be provided when you submit your survey, we will separate the user ID from the data when it arrives. From that point on, your data will be tagged with a 3 digit code number and can, in no way, be tied back to you. We are only interested in making sure each participant has completed the survey, so we will use the user ID to send reminder emails back to those participants who have not completed the survey by the deadline.

Please click the Next button to continue the survey.

Next

Done My Computer

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address A:\The Tools\CGI-BIN\TDurak-demographic_information_questionnaire_page2.htm

Demographic Information Survey

Page 2

Risks, Benefits, and Compensation:

No physical or mental risks are foreseeable. Your participation will benefit older and younger adults who need access to quality health care information. Personal satisfaction can result from knowing you have contributed to a worthwhile effort. At the end of this study, you will be compensated for your efforts.

Freedom To Withdraw:

You are free to withdraw from this study at any time without penalty.

Approval:

This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects at Virginia Polytechnic Institute and State University and by the Department of Industrial and Systems Engineering.

If you have questions at any time about the project or the procedures, you may contact the principal investigators, Tolga Durak at 961-2176, tdedu@hotmail.com or Dr. Tonya Smith-Jackson at 231-4119, smithjack@vt.edu (519-H Whittemore).

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant have been violated during the course of this project, you may contact Dr. D. M. Moore, Chair of the Institutional Review Board Research Division at 231-4991 or moored@vt.edu.

Please click the Next button to continue the survey.

Next

Done My Computer

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address A:\The Tools\CGI-BIN\TDurak-demographic_information_questionnaire_page3.htm

Demographic Information Survey

Page 3

Participant's Permission:

I have read and understand the Informed Consent and conditions of this project. I have had all questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project.

By typing your name and the date on the line below, you are communicating your consent to participate in this project.

Identification Number

Please type your 2-digit identification number into the box below. This is the number you received in a previous email. (Position the cursor in the box and begin typing.

Please click the Next button to continue the survey.

Next

Done My Computer

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address A:\The Tools\CGI-BIN\TDurak-demographic_information_questionnaire_page4.htm

Demographic Information Survey

Page 4

1. What is your gender? (Position the arrow over the circle beside the response you would like to select. Click the mouse to select.)

- Male
- Female

2. How old are you? (Position the arrow over the circle beside the response you would like to select. Click the mouse to select.)

- 18-26 Years Old
- 27-35 Years Old
- 36-44 Years Old
- 45-53 Years Old
- 54-62 Years Old
- 63-71 Years Old
- 72-80 Years Old
- 81-89 Years Old

Please click the Next button to continue the survey.

Next

Done My Computer

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address A:\The Tools\CGI-BIN\TDurak-demographic_information_questionnaire_page5.htm

Demographic Information Survey

Page 5

3. Which of the following educational levels have you achieved or are anticipating to achieve? (Position the arrow over the circle beside the response you would like to select. Click the mouse to select.)(Please choose only one answer)

- Did not complete High School
- High School diploma or G.E.D.
- Community or Junior College, Trade or Vocational school (Certificate, Associates Degree)
- Baccalaureate (B.S., B.A., etc)
- Certificate of Advanced Graduate Studies
- Masters Degree (M.A., M.S., M.Ed., etc)
- Doctoral Degree (Ph.D., Ed.D., J. D., D.D., etc)

4. Where are you located while completing this survey and where will you be located while completing the second survey? (Position the arrow over the circle beside the response you would like to select. Click the mouse to select.)

- In Assessment and Cognitive Ergonomics Lab
- At the Senior Center At Blacksburg Parks and Recreation Center
- In my Home
- Another Location.

Please click the Next button to continue the survey.

Next

Done My Computer

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address A:\The Tools\CGI-BIN\TDurak-demographic_information_questionnaire_page6.htm

Demographic Information Survey

Page 6

5. If you selected "Another Location" in question 4 above, please specify the location in the box provided below:(Position the cursor in the box and begin typing.)

6. What is the primary browser you use for web access? (Position the arrow over the circle beside the response you would like to select. Click the mouse to select.)

- Netscape Navigator/Communicator
- Microsoft Internet Explorer
- Mosaic
- America Online
- Other

7. If you have selected "Other" in question 6 above, please specify the location in the box provided below: (Position the cursor in the box and begin typing.)

Please click the Next button to continue the survey.

Next

Done My Computer

Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address A:\The Tools\CGI-BIN\TDurak-demographic_information_questionnaire_page7.htm

Demographic Information Survey

Page 7

Contact Information. Thank you for taking the time to complete The Demographic Information Questionnaire. Please feel free to contact Tolga Durak if you have any questions (tdedu@hotmail.com; 961-2176).

Please click on the submit button at the bottom of this page to submit this survey.

Submit Survey

This questionnaire was created using Perseus SurveySolutions.

Done My Computer

APPENDIX O

PARTICIPANT DEMOGRAPHIC INFORMATION FOR STUDY 1 & 2

The summary of participant demographic information for Study 1 is as follows:

1) 15 participants were male, 15 participants were female
2) 10 participants were beginner users, 10 participants were intermediate users, and 10 participants were expert users.
3) There were equal number of female and male participants in each experience level category
4) All participants are between the ages of 18 to 26.
5) All participants have achieved or anticipating to achieve a Baccalaureate degree
6) All participants completed the surveys from their house, apartments or dorm rooms
7) Out of 30 participants, 20 participants use Microsoft Internet Explorer, 8 participants use Netscape Navigator/Communicator and 2 participants use America Online as their primary browser.

The summary of participant demographic information for Study 2 is as follows:

1) 12 participants were older participants (50+), 12 participants were younger participants.
2) 12 participants were male, 12 participants were female
3) 8 participants were beginner users, 8 participants were intermediate users, and 8 participants were expert users.
4) There were equal number of female and male participants in each experience level category
5) There were equal numbers of older and younger participants in each experience level category.
6) 9 younger participants are between the ages of 18 to 26. Three younger participants are between the ages of 27-35.
7) 5 older participants are between the ages of 54-62. Three older participants are between the ages of 63-71. Four older participants are between the ages of 50-53. One older participant is between the ages of 81-89.
8) All younger participants have achieved or anticipating to achieve a Baccalaureate degree.
9) Five older participants have achieved or anticipating to achieve a High School diploma or G.E.D. Four older participants have achieved or anticipating to achieve a Community or Junior College, Trade or Vocational school certificate or associate degree. Two older participants have achieved or anticipating to achieve a Baccalaureate degree. One older participant have achieved or anticipating to achieve a Masters degree.

10) Ten younger participants completed the surveys at their house, apartments or dorm rooms. Two younger participants completed the surveys at their work.
11) Nine older participants completed the surveys at their house or apartments. Ten older participants completed the surveys at their work.
12) 8 younger participants use Microsoft Internet Explorer, 4 younger participants use Netscape Navigator/Communicator.
13) 11 older participants use Microsoft Internet Explorer, 1 older participant uses Netscape Navigator/Communicator.

APPENDIX P

THE IRB APPLICATION

Investigating the Usefulness of Age-Independent Online Needs Assessment Tools

Principal Investigators: Tolga Durak and Tonya L. Smith-Jackson, Ph.D.

Justification of Project

The popularity of online needs assessment tools is increasing because of the online application demands in general and the usefulness and efficiency of these applications. Similar to all needs assessment tools, online needs assessment tools support a systematic exploration of the current system and the expected, desired or required system to be developed. In addition, online medical assessment tools are providing patients with easy access at any time of the day or any day of the week. The service is time-efficient and can reach potential participants in isolated environments or those who have a lack of transportation. Also the tools are eliminating the problems that personal direct medical assistance might have, such as restricted mobility caused by health issues or fear of crime (Smith-Jackson & Williges, 2000).

Several studies on older users have been conducted in the literature. These studies typically focused on elderly users, Web applications, and cognitive, perceptual and motor skills related to age. Most of these studies are very case and user-specific and only cover one limited aspect of the broad area of user-centered design for older users. Although there is a deceptive amount of applications promoted to accommodate older user needs, none of these studies in the literature covered all the aspects and the considerations that are needed in order to develop the whole application. In addition,

most of the current technology design applications are often based on only younger users, preventing the developed tools from successfully targeting the design issues for older users. As a result of all this lack of empirical research in the field of online needs assessment tools designed to accommodate older users, the needs, requirements and capabilities of older users are not met completely.

This study will be conducted in order to provide information resource to eliminate the mismatch between the current applications and older users' needs, requirements, expectations and capabilities. The study will also be designed to explore the usability of an age-independent online needs assessment tool.

Method

Needs Assessment:

Participants

A total of 54 participants will be recruited from the local community. Twelve of these participants will be older adults (50 years old or over) and the remaining 42 participants will be younger adults (younger than 35 years old). Most participants will be recruited from the members of the Blacksburg Electronic Village (BEV) and from the students of Virginia Polytechnic Institute and State University. The test group sessions will be conducted in the Senior Center computer lab at the Blacksburg Parks and Recreation Center, in the Assessment Cognitive Ergonomics Lab (Whittemore Hall room 519-A) at Virginia Polytechnic Institute and State University campus or in another convenient location.

The purpose of the test group sessions is to gather information on user needs, preferences, and capabilities to validate and improve a developed online needs assessment tool.

Informed Consent: Participants will be given one of the two informed consent documents depending on the participant's age (see attached document) and the participant rights and responsibilities will be discussed at the beginning of each session. Once voluntary consent is communicated via signature, the sessions will begin.

Test Group Sessions:

A total of 54 participants (12 older adults and 42 younger adults) will complete online assessment tool prototypes and will provide feedback regarding the content of the online assessment tool, the usability of the online assessment tool and the usability of a prototype interface. Online assessments will be designed to require no more than 30-45 minutes to complete.

Once the sessions have been completed, participants will be thanked for their participation and debriefed.

RISKS AND BENEFITS

There are no factors related to this project that would place participants at more than minimal risk. No long-term emotional repercussions are foreseeable.

CONFIDENTIALITY/ANONYMITY

In order to coordinate volunteers and session schedules, participants names, phone, numbers and email addresses will be used. However, confidentiality will be upheld and participants will be assured of confidentiality. When data are collected, a three-digit code number will be used. The code number cannot be tied back to the participants' names.

Confidentiality in these sessions can only be protected to the extent that participants are free to withhold information or feedback that they are uncomfortable reporting. This participant right will be included in the informed consent document.

INFORMED CONSENT

See attached two different informed consent forms.

BIOGRAPHICAL SKETCH

Tolga Durak is a graduate student in Human Factors Engineering option in the Industrial and Systems Engineering Department. During his studies at Virginia Polytechnic Institute and State University with the Human Factors Engineering option, he conducted and was involved in research using human subjects in the areas of warning and compliance, safety information design, industrial ergonomics, cognitive ergonomics and usability testing.

Dr. Smith-Jackson is an assistant professor of Human Factors Engineering in the Industrial and Systems Engineering Department. For the past 11 years, she has conducted research using human subjects in the areas of warning and compliance, safety information design, focused attention, and cognitive ergonomics. Most of her research was conducted within Departments of Psychology. She has also conducted usability testing at various corporations such as Ericsson, IBM, and PC&InfoSystems Consulting.

She has taught Introduction to Human Factors Engineering, Human-Computer Systems, and Occupational Safety and Hazard Control. In the Intro. to HFE course, she spends a full class period on the ethical guidelines that relate to the use of human subjects in research. Also, students are given scenarios of ethical dilemmas in class

and on the first exam. They must apply the guidelines to determine the appropriate course of action.

APPENDIX Q

CHANGES AFTER STUDY 1

The following are the revisions made to qualitative and quantitative questions of Prototype B after Study 1:

1) Since participants expressed their satisfaction with the closed-ended questions, no changes will be made to these questions.
2) Even though participants didn't express their dissatisfaction with the open-ended questions, the qualitative data gathered after Study 1 was not sufficient enough to conduct content analysis or to extract guidelines. Therefore, a revision was made to qualitative questions in order to gather more information during Study 2.
3) Part I was used as a test-bed so the data gathered from this section was supplemental information. Also, the qualitative questions of Part 1 derived sufficient amount of data. Thus, no changes to the qualitative questions of Part I was necessary.
4) For the qualitative questions of Part II and Part III, the list of changes made are as follows:
    a. Instead of one qualitative question, two qualitative questions were asked at the end of each part. By this way, participants will be forced to think and respond about different aspects separately. Even though the number of questions increased, the amount of information asked still remained the same. Therefore, typing demands and the length of the questionnaire remained the same. However, the responses will be gathered in a better and more effective manner. After this revision, question 5 of Part II was changed to:

    "*Question 5:*
    *Please use the blank below to tell us what you **liked** about Part I and Part II that you just completed and why. Keep in mind that you can comment on the Online Questionnaire (Part I), Usability Questions (Part II), methods used to elicit needs, requirements and preferences in order to design healthcare information websites, questions asked and the participation process."*

    And,

    "*Question 6:*
    *Please use the blank below to tell us what you **disliked** about Part I and Part II that you just completed and why. Keep in mind that you can comment on the Online Questionnaire (Part I), Usability Questions (Part II), methods used to elicit needs, requirements and preferences in order to*

design healthcare information websites, questions asked and the participation process."

Also, Question 15 of Part III was changed to:

"*Question 15:*
If there are any concerns regarding ***a specific guideline****, please use the blank below to let us know what you* ***liked or disliked*** *about the guideline and why. Keep in mind that you can comment on the each of the features that are represented by each guideline, the success rate of each feature/guideline's implementation and importance of each feature/ guideline*"

And,

"*Question 16:*
If there are any other ***features or guidelines*** *that you think are important and should be included in this section as well as the design of the Online Questionnaire, please use the blank below to let us know. Keep in mind that each guideline consists of more than one feature*."

b. The following encouragement sentence was added right before the open-ended questions at the end of Part II and Part III:

   "*For the next two questions, please try to provide as much feedback as possible. Your input is very important and will be used to improve our process and product."*

   This sentence was added to encourage the participants to provide more information.

c. The following definition of a Healthcare Website was also added right before Part I:

   "*Healthcare Websites are websites that are designed to provide healthcare/medical information, healthcare organizations, medical products and medical advice*."

   This sentence was added in order to provide a definition of the Healthcare Website for the participants who were not knowledgeable.

APPENDIX R

Complete ANOVA results for the questions with significant differences is as follows:

Study 1

Part I:

a) Time

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: time

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	119.4666667	23.8933333	14.93	<.0001
Error	24	38.4000000	1.6000000		
Corrected Total	29	157.8666667			

R-Square	Coeff Var	Root MSE	time Mean
0.756757	17.89968	1.264911	7.066667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	112.1333333	112.1333333	70.08	<.0001
exp	2	2.0666667	1.0333333	0.65	0.5331
gender*exp	2	5.2666667	2.6333333	1.65	0.2139

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for time

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	1.6
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.9533

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	9.0000	15	1
B	5.1333	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for time

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	1.6
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	1.4127

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	7.3000	10	2
A			
A	7.2000	10	1
A			
A	6.7000	10	3

The ANOVA Procedure

Level of gender	Level of exp	N	time Mean	time Std Dev
1	1	5	9.00000000	1.41421356
1	2	5	8.80000000	1.64316767
1	3	5	9.20000000	1.30384048
2	1	5	5.40000000	1.14017543
2	2	5	5.80000000	1.30384048
2	3	5	4.20000000	0.44721360

b) Question 7, response alternative 5:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q7_5

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	48.9666667	9.7933333	2.83	0.0382
Error	24	83.2000000	3.4666667		
Corrected Total	29	132.1666667			

R-Square	Coeff Var	Root MSE	Q7_5 Mean
0.370492	27.24730	1.861899	6.833333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	36.30000000	36.30000000	10.47	0.0035
exp	2	8.86666667	4.43333333	1.28	0.2967
gender*exp	2	3.80000000	1.90000000	0.55	0.5851

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q7_5

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	3.466667
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	1.4032

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	7.9333	15	2
B	5.7333	15	1

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q7_5

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	3.466667
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	2.0794

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	7.4000	10	1
A			
A	7.0000	10	2
A			
A	6.1000	10	3

The ANOVA Procedure

Level of gender	Level of exp	N	Q7_5 Mean	Q7_5 Std Dev
1	1	5	6.60000000	1.94935887
1	2	5	5.40000000	3.04959014
1	3	5	5.20000000	1.64316767
2	1	5	8.20000000	1.09544512
2	2	5	8.60000000	0.54772256
2	3	5	7.00000000	1.87082869

Part II:

a) Question 1, response alternative 4:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q1_4

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	5.86666667	1.17333333	2.07	0.1045
Error	24	13.60000000	0.56666667		
Corrected Total	29	19.46666667			

R-Square	Coeff Var	Root MSE	Q1_4 Mean
0.301370	8.821555	0.752773	8.533333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	4.80000000	4.80000000	8.47	0.0077
exp	2	0.86666667	0.43333333	0.76	0.4765
gender*exp	2	0.20000000	0.10000000	0.18	0.8393

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q1_4

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.566667
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.5673

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.9333	15	1
B	8.1333	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q1_4

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.566667
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	0.8407

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.7000	10	1
A			
A	8.6000	10	3
A			
A	8.3000	10	2

The ANOVA Procedure

Level of gender	Level of exp	N	Q1_4 Mean	Q1_4 Std Dev
1	1	5	9.00000000	0.00000000
1	2	5	8.80000000	0.44721360
1	3	5	9.00000000	0.00000000
2	1	5	8.40000000	0.89442719
2	2	5	7.80000000	0.83666003
2	3	5	8.20000000	1.30384048

Part III:

a) Question 2, agreement scale:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q2_1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	18.66666667	3.73333333	4.23	0.0068
Error	24	21.20000000	0.88333333		
Corrected Total	29	39.86666667			

R-Square	Coeff Var	Root MSE	Q2_1 Mean
0.468227	12.15334	0.939858	7.733333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	10.80000000	10.80000000	12.23	0.0019
exp	2	0.46666667	0.23333333	0.26	0.7701
gender*exp	2	7.40000000	3.70000000	4.19	0.0275

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q2_1

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.883333
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.7083

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.3333	15	1
B	7.1333	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q2_1

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.883333
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	1.0497

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	7.9000	10	3
A			
A	7.7000	10	2
A			
A	7.6000	10	1

The ANOVA Procedure

Level of gender	Level of exp	N	Q2_1 Mean	Q2_1 Std Dev
1	1	5	8.60000000	0.54772256
1	2	5	7.60000000	1.34164079
1	3	5	8.80000000	0.44721360
2	1	5	6.60000000	1.14017543
2	2	5	7.80000000	0.83666003
2	3	5	7.00000000	1.00000000

b) Question 3, importance scale:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q3_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	12.26666667	2.45333333	4.46	0.0051
Error	24	13.20000000	0.55000000		
Corrected Total	29	25.46666667			

R-Square	Coeff Var	Root MSE	Q3_2 Mean
0.481675	9.118277	0.741620	8.133333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	8.53333333	8.53333333	15.52	0.0006
exp	2	0.46666667	0.23333333	0.42	0.6591
gender*exp	2	3.26666667	1.63333333	2.97	0.0704

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q3_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.55
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.5589

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.6667	15	1
B	7.6000	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q3_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.55
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	0.8283

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.3000	10	2
A			
A	8.1000	10	1
A			
A	8.0000	10	3

The ANOVA Procedure

Level of gender	Level of exp	N	Q3_2 Mean	Q3_2 Std Dev
1	1	5	9.00000000	0.00000000
1	2	5	8.40000000	0.89442719
1	3	5	8.60000000	0.89442719
2	1	5	7.20000000	0.44721360
2	2	5	8.20000000	0.44721360
2	3	5	7.40000000	1.14017543

c) Question 4, importance scale:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q4_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	12.26666667	2.45333333	4.46	0.0051
Error	24	13.20000000	0.55000000		
Corrected Total	29	25.46666667			

R-Square	Coeff Var	Root MSE	Q4_2 Mean
0.481675	9.427371	0.741620	7.866667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	10.80000000	10.80000000	19.64	0.0002
exp	2	0.06666667	0.03333333	0.06	0.9413
gender*exp	2	1.40000000	0.70000000	1.27	0.2983

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q4_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.55
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.5589

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.4667	15	1
B	7.2667	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q4_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.55
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	0.8283

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	7.9000	10	1
A			
A	7.9000	10	3
A			
A	7.8000	10	2

The ANOVA Procedure

Level of gender	Level of exp	N	Q4_2 Mean	Q4_2 Std Dev
1	1	5	8.80000000	0.44721360
1	2	5	8.20000000	1.09544512
1	3	5	8.40000000	0.54772256
2	1	5	7.00000000	0.70710678
2	2	5	7.40000000	0.54772256
2	3	5	7.40000000	0.89442719

d) Question 5, importance scale:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q5_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	6.96666667	1.39333333	4.18	0.0071
Error	24	8.00000000	0.33333333		
Corrected Total	29	14.96666667			

R-Square	Coeff Var	Root MSE	Q5_2 Mean
0.465479	6.900601	0.577350	8.366667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	5.63333333	5.63333333	16.90	0.0004
exp	2	0.26666667	0.13333333	0.40	0.6747
gender*exp	2	1.06666667	0.53333333	1.60	0.2227

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q5_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.333333
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.4351

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.8000	15	1
B	7.9333	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q5_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.333333
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	0.6448

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.5000	10	3
A			
A	8.3000	10	2
A			
A	8.3000	10	1

The ANOVA Procedure

Level of gender	Level of exp	N	Q5_2 Mean	Q5_2 Std Dev
1	1	5	9.00000000	0.00000000
1	2	5	8.60000000	0.89442719
1	3	5	8.80000000	0.44721360
2	1	5	7.60000000	0.54772256
2	2	5	8.00000000	0.70710678
2	3	5	8.20000000	0.44721360

e) Question 8, importance scale:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q8_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	8.66666667	1.73333333	2.81	0.0390
Error	24	14.80000000	0.61666667		
Corrected Total	29	23.46666667			

R-Square	Coeff Var	Root MSE	Q8_2 Mean
0.369318	9.655098	0.785281	8.133333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	4.80000000	4.80000000	7.78	0.0102
exp	2	3.26666667	1.63333333	2.65	0.0913
gender*exp	2	0.60000000	0.30000000	0.49	0.6207

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q8_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.616667
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.5918

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.5333	15	1
B	7.7333	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q8_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.616667
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	0.877

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.5000	10	2
A			
A	8.2000	10	3
A			
A	7.7000	10	1

The ANOVA Procedure

Level of gender	Level of exp	N	Q8_2 Mean	Q8_2 Std Dev
1	1	5	8.00000000	1.00000000
1	2	5	8.80000000	0.44721360
1	3	5	8.80000000	0.44721360
2	1	5	7.40000000	0.54772256
2	2	5	8.20000000	0.83666003
2	3	5	7.60000000	1.14017543

f) Question 9, agreement scale:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q9_1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	10.30000000	2.06000000	3.99	0.0090
Error	24	12.40000000	0.51666667		
Corrected Total	29	22.70000000			

R-Square	Coeff Var	Root MSE	Q9_1 Mean
0.453744	8.874016	0.718795	8.100000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	7.50000000	7.50000000	14.52	0.0009
exp	2	1.40000000	0.70000000	1.35	0.2770
gender*exp	2	1.40000000	0.70000000	1.35	0.2770

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q9_1

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.516667
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.5417

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.6000	15	1
B	7.6000	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q9_1

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.516667
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	0.8028

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.4000	10	2
A			
A	8.0000	10	3
A			
A	7.9000	10	1

The ANOVA Procedure

Level of gender	Level of exp	N	Q9_1 Mean	Q9_1 Std Dev
1	1	5	8.20000000	0.83666003
1	2	5	8.80000000	0.44721360
1	3	5	8.80000000	0.44721360
2	1	5	7.60000000	0.89442719
2	2	5	8.00000000	0.70710678
2	3	5	7.20000000	0.83666003

g) Question 9, importance scale:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q9_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	6.26666667	1.25333333	2.43	0.0648
Error	24	12.40000000	0.51666667		
Corrected Total	29	18.66666667			

R-Square	Coeff Var	Root MSE	Q9_2 Mean
0.335714	8.625543	0.718795	8.333333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	4.80000000	4.80000000	9.29	0.0055
exp	2	0.06666667	0.03333333	0.06	0.9377
gender*exp	2	1.40000000	0.70000000	1.35	0.2770

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q9_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.516667
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.5417

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.7333	15	1
B	7.9333	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q9_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.516667
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	0.8028

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.4000	10	2
A			
A	8.3000	10	1
A			
A	8.3000	10	3

The ANOVA Procedure

Level of gender	Level of exp	N	Q9_2 Mean	Q9_2 Std Dev
1	1	5	9.00000000	0.00000000
1	2	5	8.60000000	0.54772256
1	3	5	8.60000000	0.89442719
2	1	5	7.60000000	0.89442719
2	2	5	8.20000000	0.83666003
2	3	5	8.00000000	0.70710678

h) Question 13, importance scale:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
gender	2	1 2
exp	3	1 2 3

Number of observations 30

The ANOVA Procedure

Dependent Variable: Q13_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	5.20000000	1.04000000	2.50	0.0590
Error	24	10.00000000	0.41666667		
Corrected Total	29	15.20000000			

R-Square	Coeff Var	Root MSE	Q13_2 Mean
0.342105	7.505782	0.645497	8.600000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	4.80000000	4.80000000	11.52	0.0024
exp	2	0.20000000	0.10000000	0.24	0.7885
gender*exp	2	0.20000000	0.10000000	0.24	0.7885

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q13_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.416667
Critical Value of Studentized Range	2.91880
Minimum Significant Difference	0.4865

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	9.0000	15	1
B	8.2000	15	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q13_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	24
Error Mean Square	0.416667
Critical Value of Studentized Range	3.53170
Minimum Significant Difference	0.7209

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.7000	10	3
A			
A	8.6000	10	2
A			
A	8.5000	10	1

The ANOVA Procedure

Level of gender	Level of exp	N	Q13_2 Mean	Q13_2 Std Dev
1	1	5	9.00000000	0.00000000
1	2	5	9.00000000	0.00000000
1	3	5	9.00000000	0.00000000
2	1	5	8.00000000	1.00000000
2	2	5	8.20000000	0.83666003
2	3	5	8.40000000	0.89442719

Study 2

Part I:

a) Time

The ANOVA Procedure

Class Level Information

Class	Levels	Values
age	2	1 2
gender	2	1 2
exp	3	1 2 3

Number of observations 24

The ANOVA Procedure

Dependent Variable: time

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	220.4583333	20.0416667	16.59	<.0001
Error	12	14.5000000	1.2083333		
Corrected Total	23	234.9583333			

R-Square	Coeff Var	Root MSE	time Mean
0.938287	12.15752	1.099242	9.041667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	0.0416667	0.0416667	0.03	0.8558
exp	2	5.0833333	2.5416667	2.10	0.1648
age	1	198.3750000	198.3750000	164.17	<.0001
age*gender	1	0.0416667	0.0416667	0.03	0.8558
age*exp	2	9.2500000	4.6250000	3.83	0.0518
gender*exp	2	1.0833333	0.5416667	0.45	0.6490
age*gender*exp	2	6.5833333	3.2916667	2.72	0.1058

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for time

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	1.208333
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	0.9778

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	9.0833	12	2
A			
A	9.0000	12	1

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for time

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	1.208333
Critical Value of Studentized Range	3.77278
Minimum Significant Difference	1.4663

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	9.6250	8	1
A			
A	9.0000	8	2
A			
A	8.5000	8	3

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for time

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	1.208333
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	0.9778

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	age
A	11.9167	12	2
B	6.1667	12	1

The ANOVA Procedure

Level of age	Level of gender	N	time Mean	time Std Dev
1	1	6	6.1666667	0.98319208
1	2	6	6.1666667	0.40824829
2	1	6	11.8333333	2.04124145
2	2	6	12.0000000	1.41421356

Level of age	Level of exp	N	time Mean	time Std Dev
1	1	4	6.2500000	0.95742711
1	2	4	5.7500000	0.50000000
1	3	4	6.5000000	0.57735027
2	1	4	13.0000000	1.15470054
2	2	4	12.2500000	1.25830574
2	3	4	10.5000000	1.73205081

Level of gender	Level of exp	N	time Mean	time Std Dev
1	1	4	9.50000000	4.20317340
1	2	4	9.25000000	4.42530602
1	3	4	8.25000000	1.50000000
2	1	4	9.75000000	3.86221008
2	2	4	8.75000000	3.20156212
2	3	4	8.75000000	3.40342964

Level of age	Level of gender	Level of exp	N	time Mean	time Std Dev
1	1	1	2	6.0000000	1.41421356
1	1	2	2	5.5000000	0.70710678
1	1	3	2	7.0000000	0.00000000
1	2	1	2	6.5000000	0.70710678
1	2	2	2	6.0000000	0.00000000
1	2	3	2	6.0000000	0.00000000
2	1	1	2	13.0000000	1.41421356
2	1	2	2	13.0000000	1.41421356
2	1	3	2	9.5000000	0.70710678
2	2	1	2	13.0000000	1.41421356
2	2	2	2	11.5000000	0.70710678
2	2	3	2	11.5000000	2.12132034

b) Question 6, response alternative 7:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
age	2	1 2
gender	2	1 2
exp	3	1 2 3

Number of observations 24

The ANOVA Procedure

Dependent Variable: Q6_7

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	108.4583333	9.8598485	4.83	0.0057
Error	12	24.5000000	2.0416667		
Corrected Total	23	132.9583333			

R-Square	Coeff Var	Root MSE	Q6_7 Mean
0.815732	28.34120	1.428869	5.041667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
gender	1	7.04166667	7.04166667	3.45	0.0880
exp	2	40.08333333	20.04166667	9.82	0.0030
age	1	12.04166667	12.04166667	5.90	0.0318
age*gender	1	12.04166667	12.04166667	5.90	0.0318
age*exp	2	3.08333333	1.54166667	0.76	0.4910
gender*exp	2	23.08333333	11.54166667	5.65	0.0186
age*gender*exp	2	11.08333333	5.54166667	2.71	0.1065

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q6_7

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	2.041667
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	1.271

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	5.5833	12	1
A			
A	4.5000	12	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q6_7

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.01
Error Degrees of Freedom	12
Error Mean Square	2.041667
Critical Value of Studentized Range	3.77278
Minimum Significant Difference	1.9059

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	6.7500	8	2
A	4.7500	8	3
A			
A	3.6250	8	1

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q6_7

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.01
Error Degrees of Freedom	12
Error Mean Square	2.041667
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	1.271

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	age
A	5.7500	12	2
A	4.3333	12	1

The ANOVA Procedure

Level of age	Level of gender	N	Q6_7 Mean	Q6_7 Std Dev
1	1	6	4.16666667	2.40138849
1	2	6	4.50000000	2.34520788
2	1	6	7.00000000	1.67332005
2	2	6	4.50000000	2.50998008

Level of age	Level of exp	N	Q6_7 Mean	Q6_7 Std Dev
1	1	4	2.50000000	1.29099445
1	2	4	6.00000000	1.41421356
1	3	4	4.50000000	2.64575131
2	1	4	4.75000000	3.40342964
2	2	4	7.50000000	1.29099445
2	3	4	5.00000000	1.41421356

Level of gender	Level of exp	N	Q6_7 Mean	Q6_7 Std Dev
1	1	4	5.25000000	2.98607881
1	2	4	7.50000000	1.29099445
1	3	4	4.00000000	1.82574186
2	1	4	2.00000000	0.81649658
2	2	4	6.00000000	1.41421356
2	3	4	5.50000000	2.08166600

Level of age	Level of gender	Level of exp	N	Q6_7 Mean	Q6_7 Std Dev
1	1	1	2	3.00000000	1.41421356
1	1	2	2	7.00000000	1.41421356
1	1	3	2	2.50000000	0.70710678
1	2	1	2	2.00000000	1.41421356
1	2	2	2	5.00000000	0.00000000
1	2	3	2	6.50000000	2.12132034
2	1	1	2	7.50000000	2.12132034
2	1	2	2	8.00000000	1.41421356
2	1	3	2	5.50000000	0.70710678
2	2	1	2	2.00000000	0.00000000
2	2	2	2	7.00000000	1.41421356
2	2	3	2	4.50000000	2.12132034

Study 2

Part II:

a) Time:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
age	2	1 2
gender	2	1 2
exp	3	1 2 3

Number of observations 24

The ANOVA Procedure

Dependent Variable: time

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	90.83333333	8.25757576	14.16	<.0001
Error	12	7.00000000	0.58333333		
Corrected Total	23	97.83333333			

R-Square	Coeff Var	Root MSE	time Mean
0.928450	11.90279	0.763763	6.416667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
age	1	80.66666667	80.66666667	138.29	<.0001
gender	1	0.16666667	0.16666667	0.29	0.6027
exp	2	0.58333333	0.29166667	0.50	0.6186
gender*exp	2	0.58333333	0.29166667	0.50	0.6186
age*gender	1	0.66666667	0.66666667	1.14	0.3061
age*exp	2	5.08333333	2.54166667	4.36	0.0378
age*gender*exp	2	3.08333333	1.54166667	2.64	0.1119

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for time

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.583333
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	0.6794

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	age
A	8.2500	12	2
B	4.5833	12	1

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for time

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.583333
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	0.6794

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	6.5000	12	2
A			
A	6.3333	12	1

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for time

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.583333
Critical Value of Studentized Range	3.77278
Minimum Significant Difference	1.0188

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	6.6250	8	1
A			
A	6.3750	8	2
A			
A	6.2500	8	3

The ANOVA Procedure

Level of gender	Level of exp	N	time Mean	time Std Dev
1	1	4	6.75000000	2.75378527
1	2	4	6.25000000	1.50000000
1	3	4	6.00000000	1.82574186
2	1	4	6.50000000	2.88675135
2	2	4	6.50000000	3.00000000
2	3	4	6.50000000	1.29099445

Level of age	Level of gender	N	time Mean	time Std Dev
1	1	6	4.66666667	0.51639778
1	2	6	4.50000000	1.04880885
2	1	6	8.00000000	1.09544512
2	2	6	8.50000000	0.83666003

Level of age	Level of exp	N	time Mean	time Std Dev
1	1	4	4.25000000	0.50000000
1	2	4	4.50000000	1.00000000
1	3	4	5.00000000	0.81649658
2	1	4	9.00000000	0.81649658
2	2	4	8.25000000	0.95742711
2	3	4	7.50000000	0.57735027

Level of age	Level of gender	Level of exp	N	time Mean	time Std Dev
1	1	1	2	4.50000000	0.70710678
1	1	2	2	5.00000000	0.00000000
1	1	3	2	4.50000000	0.70710678
1	2	1	2	4.00000000	0.00000000
1	2	2	2	4.00000000	1.41421356
1	2	3	2	5.50000000	0.70710678
2	1	1	2	9.00000000	1.41421356
2	1	2	2	7.50000000	0.70710678
2	1	3	2	7.50000000	0.70710678
2	2	1	2	9.00000000	0.00000000
2	2	2	2	9.00000000	0.00000000
2	2	3	2	7.50000000	0.70710678

b) Question 2, response alternative 4:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
age	2	1 2
gender	2	1 2
exp	3	1 2 3

Number of observations 24

The ANOVA Procedure

Dependent Variable: Q2_4

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	27.50000000	2.50000000	6.00	0.0022
Error	12	5.00000000	0.41666667		
Corrected Total	23	32.50000000			

R-Square	Coeff Var	Root MSE	Q2_4 Mean
0.846154	7.824209	0.645497	8.250000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
age	1	0.16666667	0.16666667	0.40	0.5390
gender	1	4.16666667	4.16666667	10.00	0.0082
exp	2	9.75000000	4.87500000	11.70	0.0015
gender*exp	2	11.08333333	5.54166667	13.30	0.0009
age*gender	1	1.50000000	1.50000000	3.60	0.0821
age*exp	2	0.08333333	0.04166667	0.10	0.9056
age*gender*exp	2	0.75000000	0.37500000	0.90	0.4323

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q2_4

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.416667
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	0.5742

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	age
A	8.3333	12	2
A			
A	8.1667	12	1

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q2_4

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.416667
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	0.5742

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.6667	12	1
B	7.8333	12	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q2_4

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.416667
Critical Value of Studentized Range	3.77278
Minimum Significant Difference	0.861

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.8750	8	3
A			
A	8.5000	8	1
B	7.3750	8	2

The ANOVA Procedure

Level of gender	Level of exp	N	Q2_4 Mean	Q2_4 Std Dev
1	1	4	8.50000000	1.00000000
1	2	4	8.75000000	0.50000000
1	3	4	8.75000000	0.50000000
2	1	4	8.50000000	0.57735027
2	2	4	6.00000000	0.81649658
2	3	4	9.00000000	0.00000000

Level of age	Level of gender	N	Q2_4 Mean	Q2_4 Std Dev
1	1	6	8.33333333	0.81649658
1	2	6	8.00000000	1.54919334
2	1	6	9.00000000	0.00000000
2	2	6	7.66666667	1.50554531

Level of age	Level of exp	N	Q2_4 Mean	Q2_4 Std Dev
1	1	4	8.50000000	1.00000000
1	2	4	7.25000000	1.50000000
1	3	4	8.75000000	0.50000000
2	1	4	8.50000000	0.57735027
2	2	4	7.50000000	1.91485422
2	3	4	9.00000000	0.00000000

Level of age	Level of gender	Level of exp	N	Q2_4 Mean	Q2_4 Std Dev
1	1	1	2	8.00000000	1.41421356
1	1	2	2	8.50000000	0.70710678
1	1	3	2	8.50000000	0.70710678
1	2	1	2	9.00000000	0.00000000
1	2	2	2	6.00000000	0.00000000
1	2	3	2	9.00000000	0.00000000
2	1	1	2	9.00000000	0.00000000
2	1	2	2	9.00000000	0.00000000
2	1	3	2	9.00000000	0.00000000
2	2	1	2	8.00000000	0.00000000
2	2	2	2	6.00000000	1.41421356
2	2	3	2	9.00000000	0.00000000

c) Question 3, response alternative 2:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
age	2	1 2
gender	2	1 2
exp	3	1 2 3

Number of observations 24

The ANOVA Procedure

Dependent Variable: Q3_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	24.50000000	2.22727273	4.45	0.0080
Error	12	6.00000000	0.50000000		
Corrected Total	23	30.50000000			

R-Square	Coeff Var	Root MSE	Q3_2 Mean
0.803279	9.123958	0.707107	7.750000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
age	1	1.50000000	1.50000000	3.00	0.1089
gender	1	2.66666667	2.66666667	5.33	0.0395
exp	2	3.25000000	1.62500000	3.25	0.0745
gender*exp	2	8.58333333	4.29166667	8.58	0.0049
age*gender	1	6.00000000	6.00000000	12.00	0.0047
age*exp	2	2.25000000	1.12500000	2.25	0.1480
age*gender*exp	2	0.25000000	0.12500000	0.25	0.7828

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q3_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.5
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	0.629

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	age
A	8.0000	12	2
A			
A	7.5000	12	1

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q3_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.5
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	0.629

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	8.0833	12	1
B	7.4167	12	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q3_2

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	0.5
Critical Value of Studentized Range	3.77278
Minimum Significant Difference	0.9432

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	8.2500	8	3
A			
A	7.6250	8	2
A			
A	7.3750	8	1

The ANOVA Procedure

Level of gender	Level of exp	N	Q3_2 Mean	Q3_2 Std Dev
1	1	4	8.00000000	1.15470054
1	2	4	8.50000000	0.57735027
1	3	4	7.75000000	1.25830574
2	1	4	6.75000000	0.50000000
2	2	4	6.75000000	1.25830574
2	3	4	8.75000000	0.50000000

Level of age	Level of gender	N	Q3_2 Mean	Q3_2 Std Dev
1	1	6	7.33333333	0.81649658
1	2	6	7.66666667	1.21106014
2	1	6	8.83333333	0.40824829
2	2	6	7.16666667	1.32916014

Level of age	Level of exp	N	Q3_2 Mean	Q3_2 Std Dev
1	1	4	6.75000000	0.50000000
1	2	4	7.75000000	0.50000000
1	3	4	8.00000000	1.41421356
2	1	4	8.00000000	1.15470054
2	2	4	7.50000000	1.91485422
2	3	4	8.50000000	0.57735027

Level of age	Level of gender	Level of exp	N	Q3_2 Mean	Q3_2 Std Dev
1	1	1	2	7.00000000	0.00000000
1	1	2	2	8.00000000	0.00000000
1	1	3	2	7.00000000	1.41421356
1	2	1	2	6.50000000	0.70710678
1	2	2	2	7.50000000	0.70710678
1	2	3	2	9.00000000	0.00000000
2	1	1	2	9.00000000	0.00000000
2	1	2	2	9.00000000	0.00000000
2	1	3	2	8.50000000	0.70710678
2	2	1	2	7.00000000	0.00000000
2	2	2	2	6.00000000	1.41421356
2	2	3	2	8.50000000	0.70710678

d) Question 3, response alternative 3:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
age	2	1 2
gender	2	1 2
exp	3	1 2 3

Number of observations 24

The ANOVA Procedure

Dependent Variable: Q3_3

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	110.4583333	10.0416667	3.95	0.0130
Error	12	30.5000000	2.5416667		
Corrected Total	23	140.9583333			

R-Square	Coeff Var	Root MSE	Q3_3 Mean
0.783624	27.92865	1.594261	5.708333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
age	1	30.37500000	30.37500000	11.95	0.0047
gender	1	30.37500000	30.37500000	11.95	0.0047
exp	2	1.58333333	0.79166667	0.31	0.7381
gender*exp	2	9.75000000	4.87500000	1.92	0.1893
age*gender	1	30.37500000	30.37500000	11.95	0.0047
age*exp	2	3.25000000	1.62500000	0.64	0.5447
age*gender*exp	2	4.75000000	2.37500000	0.93	0.4196

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q3_3

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	2.541667
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	1.4181

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	age
A	6.8333	12	1
B	4.5833	12	2

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q3_3

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	2.541667
Critical Value of Studentized Range	3.08132
Minimum Significant Difference	1.4181

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	gender
A	6.8333	12	2
B	4.5833	12	1

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Q3_3

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	2.541667
Critical Value of Studentized Range	3.77278
Minimum Significant Difference	2.1265

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	exp
A	6.0000	8	2
A			
A	5.7500	8	3
A			
A	5.3750	8	1

The ANOVA Procedure

Level of gender	Level of exp	N	Q3_3 Mean	Q3_3 Std Dev
1	1	4	4.00000000	3.55902608
1	2	4	4.25000000	2.75378527
1	3	4	5.50000000	3.00000000
2	1	4	6.75000000	0.50000000
2	2	4	7.75000000	1.25830574
2	3	4	6.00000000	1.41421356

Level of age	Level of gender	N	Q3_3 Mean	Q3_3 Std Dev
1	1	6	6.83333333	0.75277265
1	2	6	6.83333333	0.75277265
2	1	6	2.33333333	2.42212028
2	2	6	6.83333333	1.72240142

Level of age	Level of exp	N	Q3_3 Mean	Q3_3 Std Dev
1	1	4	7.00000000	0.81649658
1	2	4	6.75000000	0.95742711
1	3	4	6.75000000	0.50000000
2	1	4	3.75000000	3.20156212
2	2	4	5.25000000	3.86221008
2	3	4	4.75000000	2.87228132

Level of age	Level of gender	Level of exp	N	Q3_3 Mean	Q3_3 Std Dev
1	1	1	2	7.00000000	1.41421356
1	1	2	2	6.50000000	0.70710678
1	1	3	2	7.00000000	0.00000000
1	2	1	2	7.00000000	0.00000000
1	2	2	2	7.00000000	1.41421356
1	2	3	2	6.50000000	0.70710678
2	1	1	2	1.00000000	0.00000000
2	1	2	2	2.00000000	1.41421356
2	1	3	2	4.00000000	4.24264069
2	2	1	2	6.50000000	0.70710678
2	2	2	2	8.50000000	0.70710678
2	2	3	2	5.50000000	2.12132034

APPENDIX S

SUPPLEMENTAL RESULTS FOR STUDY II:

a) Close-Ended Questions with check-box alternatives:

In Part I, a total of five close-ended questions (Questions 1,5,11,12, and 13) were used. For each of these questions and their response alternatives, a separate 2*2*3 Chi-Square Test was performed. Chi-Square Tests were performed for experience level categories (exp) and test groups 2 and 3 (T.G.) as well as gender, due to the equal number of male and female participants. In addition, for several response alternatives, Chi-Square test was identified as an invalid test due to having fewer than 5 elements per a response cell. For these response alternatives, separate Mantel-Haenszel Chi-Square tests were performed in order to determine the significant differences.

The first question, which was identified as Question 1 in Part I, was used to elicit the frequency for each Internet use motive. For this question, all of the participants chose one or more of the response alternatives (as shown in Table 18) for their Internet use motive(s). The results provided frequency for each motive (as shown in Table 2). However, gender, age, experience level categories were found to have no significant effect ($p > .05$) on Internet use motives represented with the response alternatives.

Table 18: The Internet Use

Response Alternatives (R. A.)	I USE the internet for this item	I DO NOT USE the internet for this item
Check, send, forward, write, reply to an email	96%	4%
Participating in online Chats	70%	30%
Searching for information	100%	0%
Downloading new software and updates from the Internet	70%	30%
Checking stock quotes, news, weather, or other "daily" information	66%	33%
Entertainment: Playing games and listening to music	70%	30%

The second question was identified as Question 5 in Part I. This question asked the participants to select the content items that are important and should be included in a healthcare Web site. All twenty-four of the participants chose one or more of the response alternatives (as shown in Table 19) as important content item(s). The results provided the frequency for each content item. However, gender, age and experience level categories were found to have no significant effect ($p > .05$) on important content items.

Table 19: Important Content Items

Response Alternatives	IMPORTANT to include	NOT IMPORTANT to include
Where to find other links and information sources such as books and organizations that provide additional information	92%	8%
Definitions of diseases	88%	12%
Definitions of medical terms	92%	8%
Drug uses, interactions (with food and herbal remedies), and side effects	80%	20%
Health care benefits programs	84%	16%
Frequently asked questions	96%	4%
Herbal remedies	84%	16%
Information about health care plans and associated physicians	75%	25%
Prevention information such as nutrition and exercise information	96%	4%
Directory system with hospital numbers, local ambulance numbers, maps to local hospitals and clinics	80%	20%
Directory/bulletin board of local events, educational programs/workshops, health-related meetings	71%	29%

The third and the forth questions, which were identified as Questions 11 and 12 in Part I, was used to elicit information on FAQ (Frequently Asked Question) section usefulness. The results showed that 88% of the participants (21 out of 24 participants) have used and the remaining 12% of the participants have not used the FAQ section. In addition, 63% of the participants (15 out of 24 participants) thought FAQ section would

be useful to them, 16% of the participants did not think that the FAQ section would be useful to them and the remaining 21% did not have any opinion. However, gender, age and experience level categories were found to have no significant effect on ($p > .05$) on FAQ section usefulness.

The last question, which was identified as Question 13 of Part I, was used to elicit information on FAQ format preference. This question asked the participants to select from one of the three response alternatives and indicate their preferred FAQ format. All twenty-four of the participants selected one FAQ format as their preference (as shown in Table 20). The results provided the frequency for each FAQ format. However, gender, age and experience level categories were found to have no significant effect ($p > .05$) on the FAQ formats represented with response alternatives.

Table 20: FAQ Format

Response Alternative	I WOULD prefer to have	I WOULD NOT prefer to have
A FAQ Section that allows me to ask any question or any topic	67%	33%
A FAQ Section that restricts questions to just a few topics (such as medications or disease symptoms).	29%	71%
A FAQ Section that restricts questions to one topic only (such as medications or disease symptoms).	4%	96%

Subsequent to data analysis and interpretation, guideline extraction methodology was applied to the results of these five questions. However, neither one of these questions or their response alternatives met the guideline extraction criteria (see Section 3.1.4 for more details on the criteria). Therefore, no item was excluded from the response alternatives and no guideline was developed.

<u>b) Close-Ended Questions with radio-button alternatives (9-point Likert-Type scale questions):</u>

The 9-point Likert-Type scale questions were used to have participants rate the importance of various healthcare contents and features that were presented with response alternatives. The mean ratings for each item are provided in Table 21.

Table 21: Healthcare Contents and Features

Item	Mean Ratings	SD
Symptoms and treatments for diseases	**7.38**	**0.21**
Pharmaceutical information: Drug information, interactions, side effects	**7.33**	**0.26**
Definitions of diseases and medical terms	**7.25**	**0.23**
Health care plans and physicians associated with each health plan (regional, local).	**7.21**	**0.28**
Knowing the sponsor(s) of a Web site you are using (for instance, insurance company sponsors, pharmaceutical sponsors)?	**7.16**	**0.30**
Search engine with an easy filter system	**7.13**	**0.26**
Exercise/nutrition/prevention information	**6.96**	**0.34**
Educational programs, events, health-related meetings in local area	**6.92**	**0.55**
Directory system with hospital numbers, etc., email addresses of physicians, phone numbers of physicians, maps	**6.83**	**0.98**
Self-Care services such as wellness tests, diet and fitness calculators	**6.63**	**1.12**
Information on which physicians in local area accept Medicare	**6.58**	**1.32**
Where to find other information sources such as other links, books, organizations	**6.58**	**1.22**
Health care benefit programs: Caregivers, assisted living, hospices, durable medical equipment, home health care	**6.58**	**1.32**
Emergency actions or special actions in case of death occurring in home	**6.33**	**1.42**
Sites to purchase medications	**6.33**	**1.12**
Opportunities to give feedback about the design for the Web site content	**6.13**	**1.32**
Hospice services and bereavement counseling	**6.08**	**1.31**
Research news, special reports, and clinical trials	**5.88**	**1.33**
Frequently asked questions option	**5.83**	**1.53**
Herbal remedies	**5.04**	**1.42**
Chats, aging healthy chats, bulletin board system (or discussion groups)	**4.92**	**1.42**

Subsequent to calculating the mean ratings, guideline extraction methodology was applied to the results of these four questions. However, neither one of the response alternatives met the guideline extraction criteria (see Section 3.1.4 for more details on the criteria). Therefore, no guideline was developed.

In addition to the guideline extraction, 2*2*3 ANOVA was performed for each of the response alternatives for completion times in order to determine the significant differences. Each ANOVA test was performed for experience level categories as well as gender and Test Groups, due to the equal number of male and female participants (see Appendix R for ANOVA results). The results of these analyses indicated two significantly different results. In addition to completion time, only one response alternative showed significantly different results.

For completion time, the following significant effect was found:

- There was a significant difference between the mean completion time for older and younger participant groups ($F\ (1,12) = 164.17,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean completion time for older participants ($M = 11.90,\ SD = 1.67$) was significantly higher than the mean completion time for younger participants ($M = 6.17,\ SD = 0.72$).

For the content item "Frequently asked questions option" (response alternative 7 of question 6), the following significant effect was found:

- There was a significant difference between the mean rating value for intermediate and the other two experience level categories ($F\ (2,12) = 9.82,\ p < .05$). Tukey's Studentized Range (HSD) test indicated that the mean rating value for intermediate experience level ($M = 6.75,\ SD = 1.49$) was significantly higher than the mean rating value for beginner experience level ($M = 3.63,\ SD = 2.67$) and expert experience level ($M = 4.75,\ SD = 1.98$). Intermediate participants expressed a significantly stronger preference towards including Frequently Asked Questions option in a healthcare website.

The Cronbach coefficient alpha value was .92 and the range of item-to-total correlation was .07 – .89. The correlation coefficients among the 21 response alternatives were all positive and statistically significant at the .05 level. The correlation coefficients ranged from a low .07 to a high .89, indicating the degree of the relationships between the response alternatives. The lowest reliable variance of the total scores was at least 92.1%. Each response alternative had adequate internal consistency and there was no need to further improve the reliability of these questions and their response alternatives.

c) Open-ended Questions, comments and suggestions sections:

In Part I, a total of four open-ended questions (Questions 2,3,4, and10) were used. The first three open-ended questions (Questions 2,3, and 4) were directed to elicit information on types of problems participants experienced when using computers and the W.W.W. The last open-ended question (Question 10) was directed to elicit any additional comments participants had regarding healthcare web site content. For each of these questions, a separate content analysis was performed. Upon the completion of content analysis, for each question, final response categories and their response frequencies were identified and guidelines extraction methodology was applied.

The first question, which was identified as Question 2, was a follow-up question that was used to explore additional ways the participants used the Internet for. For this question, two one final response categories were identified. Even though 8 participants expressed that they used the Internet for “Downloading/Checking class notes, and other class material” and 2 participants expressed that they used Internet for “Limited shopping; visiting military web sites”, these items were not relevant to the content of the

Online Tool. Thus, the items were not developed into a response alternative and were not included in the list of the Internet Use Motives.

The second and third questions, which were identified as Question 3 and 4, were used to elicit the types of computer and Internet problems participants experienced. For computer problems, ten final response categories and for Internet problems, nine final response categories were identified (as shown in Table 23 and Table 24 respectively). Several of these final response categories supported existing guidelines that expressed the importance of simplicity, clarity, ease of use, information reliability and consistency. But none of the categories included new content or application area that could be developed into a new design guideline.

Table 23: Summary of Computer Problems

Problem		Frequency
Younger Users	Not being able to find certain procedures or commands	2
	Slow Internet connection	1
	Slow Processor	8
	Computer Freezing/Locking up (Not enough memory)	6
	Computer not having enough memory to run certain programs	2
	Not being knowledgeable about computer terms	1
	Not being able to find certain procedures or commands	2
Older Users	Microsoft blue screen (system crash)	1
	Slow Processor	3
	Computer not having enough memory to run certain programs	4
	Computer Freezing/Locking up (Not enough memory)	3
	Unstable internet connection	3
	Not being knowledgeable about computers or their procedures	6
	Not knowing how to use features such as auto-save	6

Table 24: Summary of Internet Problems

Problem		Frequency
Younger Users	Websites not working or not being able to reach a site or its pages (coding or maintenance problems)	3
	Not being able to find the sources of some online information or not being sure whether or not a certain information is reliable	4
	Slow opening of web pages	6
	Slow Internet connection or not being able to maintain fast Internet connection	6
	Not being able to find desired information while conducting searches	1
	Not being able to set or correct Internet settings	2
	Freezing the computer due to running too many programs	4
Older Users	Not being able to find desired information while conducting searches	6
	Spam and pop-up adds freezing the computer	6
	Slow Internet connection or not being able to maintain fast Internet connection	5
	Websites not working or not being able to reach a site or its pages (coding or maintenance problems)	4
	Not being able to find desired information while conducting searches	6

The last question, which was identified as Question 10, was used to elicit any additional content participants would like to have on a healthcare information web site. For this question, nine final response categories were identified (as shown in Table 25). Three of these categories included new content items that were not included in the current ONAT. However, neither one of these additional content categories exceeded the overall score 0.4 (see Section 3.1.4 for more details on calculating the score). Therefore, no new content item was included in the next design iteration.

Table 25: Summary of Additional Content

Problem		Frequency
Younger Users	Information about sexually transmitted diseases	1
	Information on alternative treatments, meditation and exercise	1
	Treatment comparisons	2
	Information about AIDS\HIV	1
	Information source, its validity	1
Older Users	Books, other websites, medications and side effects	1
	Physician websites	1
	Symptoms of an illness	1
	Discussion Boards where questions to real Doctors can be asked and answered.	1
	Alternative therapies and any controversies associated with treatment	2
	Website sponsor(s), dates/times of its information and any corroborating info	1