TYPING OVER AUTOCOMPLETE: COGNITIVE LOAD IN WEBSITE USE BY OLDER ADULTS

Dr. Wendy Doube
Swinburne University of Technology
John Street Hawthorn VIC 3122
wdoube@swin.edu.au
+62 3 9214 5977

Jeanie Beh
Hawthorn Community Education Centre
Hawthorn Municipal Offices
360 Burwood Road, Hawthorn Vic 3122
Jeanie_beh@yahoo.com

ABSTRACT
This paper describes an investigation into the interaction of cognitive processing with visual and motor skills during website use by older adults. Multiple sources of data obtained from 28 older adults and 18 younger adults included self-report, task completion, observation, screen capture and video recordings, providing insight into the overall interaction of the different modalities and cognition from the perspective of cognitive load. In an online air ticketing search task, older adults took significantly longer and made significantly more mistakes than the younger control group. They appeared ignorant in the use of some screen objects and were less likely to scroll below the fold line to check results. Their longer task times were composed of a significantly higher number of gazes at the screen. The duration of their mean screen gaze was significantly longer than that of the control group. Unlike their younger counterparts, they did not glance at the screen and only checked their typing when they had completed a form field. Accordingly, they typed into autocomplete combo boxes, ignoring preset options. Avoidance of preset options caused three out of eight prominent air ticketing websites to return incorrect search results, with either higher fares or fewer flights. Screen capture software enabled the investigation to establish that the source of the errors was faulty software and not user error. Analysis of the results indicated that the longer task times, but not necessarily the number of errors, could be reduced by automating repeatable user interface actions with experience and practice.

KEYWORDS
Website, Usability, Ageing, Software testing, Graphical user interfaces, Interaction styles

1. INTRODUCTION

Ageing can be accompanied by progressive deterioration in visual, auditory, motor and cognitive functioning (Chadwick-Dias, Tedesco, & Tullis, 2004; Czaja & Lee, 2007; Money, Fernando, Lines, & Elliman, 2010; Riemersma, 1999; Whalley, 2002). Use of personal computers depends to some degree on all of these functions. Guidelines for the design and construction of user-friendly websites have been freely available since the mid to late 1990’s (Nielsen, 1994). Subsequent research has strengthened many of the original usability concepts, in particular guidelines for accessibility for marginalized website users such as older adults (Kurniawan & Zaphiris, 2005; Morrel et al., 2002; Nielsen & Norman, 2002) but these guidelines are not necessarily followed by web designers. Furthermore, implementation of guidelines may require significant judgment because the foundation research may be insufficiently generalisable (Money, et al., 2010) or because conflicting
principles may need to be balanced for different user requirements (Tullis, Tranquada, & Siegel, 2011).

1.1. Cognitive load in website usability for older adults

Cognitive deterioration with age can be observed in declines in working, spatial, situational and episodic memory and associated difficulty in learning new information and longer processing and response times (Arch, Abou-Zara, & Henry, 2008; Czaja & Lee, 2007; Priest, Nayak, & Stuart-Hamilton, 2007). Websites are easier to use if their user interfaces are designed to reduce demand on short-term memories, thus freeing cognitive resources to process the content of the site (Albers, 2011; Czaja & Lee, 2007; Hinckley, 2007). In particular, the germaine load imposed by the interface should be minimal, to free sufficient cognitive resources to process the intrinsic load of the website content and activities (Sweller, 2005).

Cognitive load increases with the number of elements in a task or object (Sweller, 2005). Working memory can accommodate on a few elements of information for a very short time. With repetition, those elements are “chunked” into fewer elements until eventually a task becomes automatic, placing minimal demands on working memory. Consistent layout and navigation throughout a site (Morrel, et al., 2002) would encourage chunking. Experienced web users to some extent have automated the processes in manipulating input devices and screen objects. A touch typist will have automated the processes in typing (Hinckley, 2007), freeing even more cognitive resources. Lower performance by older adults during web activities could be associated as much with lack of website experience as with cognitive decline (Chadwick-Dias, et al., 2004; Priest, et al., 2007).

According to the Spatial Contiguity Principle of Cognitive Theory of Multimedia Learning, visual elements of a page or a screen are easier to relate to each other if they are spatially close to each other rather than distant (Mayer, 2005). For example, a diagram with embedded labels is easier to learn than a diagram that is described by an adjacent block of text. Likewise, according to the Temporal Contiguity Principle, elements in visual and other media are easier to relate to each other if they are synchronous or close together in time (Mayer, 2005). Application of these principles in website design can reduce demands on working memory. Spatial contiguity cannot be maintained when there is too much information to fit on a single screen. Information that can be viewed only by scrolling below the fold line is less likely to be viewed (Morrel, et al., 2002). Information that is placed on another window might not be viewed at all (Czaja & Lee, 2007). In both cases, by the time the next set of information appears, it is no longer temporally or spatially contiguous with the related information previously viewed, heard or touched. Furthermore, for those who do not touch type, the distance from keyboard to screen, and the time required to locate the typed information on the screen, could also impede spatial and temporal contiguity.

Screen objects such as radio buttons, check boxes, list boxes, combo boxes, drop-down lists and menus can reduce demand on working memory by supporting recognition rather than recall (Kurniawan & Zaphiris, 2005) during selection operations while also saving space on a screen. In contrast, text boxes and sometimes combo boxes require users to...
recall information in search and selection operations. In at least two studies, user preference and performance was higher for text boxes than for selection controls for entry of dates, but in other studies, selection tasks were slower and less accurate for text boxes than for check boxes and multiple select list boxes (Tullis, et al., 2011). Radio buttons, check boxes, and list boxes display all possible options, occupying more screen space than do drop-down lists and combo boxes, which present like a single line text box, displaying one option only until the inside of the box is clicked on, after which the multiple options of the drop-down list are displayed. Depending on the number of options and the amount of space available, the number of options could exceed the drop-down space available and the user might need to scroll up or down the list to locate the right option. Drop down lists that hide one or more options would have less spatial and temporal contiguity than those that display all options on a screen without any user interaction. Furthermore, drop-down lists might hide salient information layered beneath them on the screen. The autocomplete feature in text boxes, combo boxes and address bars could also save space and support recall but very little literature can be found on this topic beyond development and security techniques.

Declines in vision include increased sensitivity to glare and decreased visual acuity, color discrimination and contrast sensitivity (Czaja & Lee, 2007; Rogers, Stronge, & Fisk, 2006). Strategies to design for declines in vision, such as increased font size, clearer grouping of information, and higher contrast between background and foreground, (Kurniawan & Zaphiris, 2005) place even greater demands on the limited screen space available.

With ageing, motor skills can also decline. Slower and more variable movements can impede the fine motor co-ordination required for control of mouse and keyboard (Bailey, 2001; Czaja & Lee, 2007; Hanson & Crayne, 2005; Rogers, et al., 2006). Strategies to help users with declining fine motor skills include large targets, avoidance of the double click option and minimal dragging (Czaja & Lee, 2007; Kurniawan & Zaphiris, 2005). Scrollbar controls and the options in drop-down boxes and lists can be too small for easy manipulation by ageing users. Furthermore, it is possible for the option in a box to be changed, possibly without the user’s knowledge, if the user does not click outside the box before attempting to scroll down the screen (Healey, 2007).

Although research in web usability for older users agrees that demands on working, spatial and procedural memory should be minimal (Czaja & Lee, 2007; Morrel, et al., 2002) the results are sometimes contradictory. Although older users can have spatial and navigation problems with multiple windows (Czaja & Lee, 2007), small amounts of information on the screen are advisable, and scrolling should be avoided (Morrel, et al., 2002). Although drop-down lists and menus can be difficult to manipulate, users with average and reduced fine motor co-ordination can still prefer them to alternative options which use screen space less effectively and require scrolling, placing demands on working memory, as well as being difficult to manipulate in their own right (Kurniawan & Zaphiris, 2005; Tullis, et al., 2011). In a 1992 study, word-processing performance by older adults was faster and contained fewer errors with drop-down lists than with on-screen menus and function keys (Czaja & Lee, 2007). Alternatively, mutually exclusive selection and task completion can be faster with radio buttons than with list boxes and drop-down lists (Tullis, et al., 2011). Nonmutually exclusive selection can be faster with checkboxes than
with multiple combo boxes or text boxes. In web surveys, higher item non-responses and longer response times were associated with drop-down lists than with radio-buttons (Healey, 2007).

1.2. Investigating website usability for multiple modalities and cognition

Many processes and tasks in website use can appear elemental to the user, whereas they are actually complex and compound, requiring visual, motor, and cognitive skills (Hinckley, 2007). Human Computer Interaction can be considered perceptual–motor interaction, and experimental conditions can be more effective if cognition and multiple modalities are considered together (Welsh, Weeks, Chua, & Goodman, 2007). Consequently multiple data collection methods are recommended. Cognitive complexity can be associated with longer time on task. Likewise, longer gaze duration on a section of the screen is likely to indicate cognitive complexity, but it can also indicate attention and interest (Blackler, Maher, & Popovic, 2010; Iqbal & Bailey, 2004; Volk, Pappas, & Wong, 2011). Web logs and screen capture software can record the movements of the mouse and keyboard with some accuracy (Kinley & Tjondronegro, 2010; Volk, et al., 2011) and consequently assist in the identification of tasks and activities that the user performs incorrectly. Although mouse and keyboard movements can be reasonably well correlated with eye movements (Cooke, 2006) time lags when the user is gazing at the screen or distracted rather than confused cannot be easily distinguished by screen capture software. In contrast, eye tracking software can indicate areas of attention and complexity and the processes leading to errors (Bojko, 2006). However, eye tracking software can be intrusive and can force the user to maintain an unnaturally still position (Zhu & Ji, 2007). Changes in head pose can possibly indicate emotion and other activities like viewing the keyboard can be missed unless other measures like video cameras are used (Asteriadis, Tzouveli, Karpouzis, & Kollias, 2008).

2. THE STUDY

The study documented in this paper investigated website usability for older adults by simultaneously exploring the interaction of visual and motor skills with cognition. Data was gathered by several means, including screen capture software, camcorder recordings, live observation, questionnaires and tests (Volk, et al., 2011). In this study older adults were classified as aged 55 or more (Chadwick-Dias, et al., 2004; Priest, et al., 2007). They were studied during web search and form completion, common and compelling web activities (Jarrett & Gaffney, 2009; Money, et al., 2010; Priest, et al., 2007).

The study was exploratory and observational. The design was iterative, consisting of three stages with further stages planned. The second and third stages were designed to deepen understanding of factors observed in the previous stage. The first and third stages were quasi experimental between groups studies with a post-test.

- In the first stage an experimental group of nine older adults aged over 55 and a control group of seven younger adults aged under 25 were given a web search task to be
performed in a computer laboratory. Participant behavior was observed and noted while on-screen actions were captured and stored on file.

- The second stage of the study was a comparative evaluation of eight different air ticketing web sites to investigate prevalence of the usability findings raised in the first stage.
- The third stage of the study repeated the first stage while also measuring gaze duration. Eighteen older adults and eight younger adults participated, bringing the total sample size to 42.

2.1. Experimental design

2.1.1 Preliminary research

A community of website users taking classes from a non-profit local organization (Hawthorn Community Education Centre) was available to participate. The organization operates from a computer laboratory in a university building. Before commencing the study, two preliminary steps were taken to ensure that procedures were suitable for the participants. Firstly, twelve older adults were surveyed in order to find a task with relevance to everyday life for their age group. The survey suggested that domestic air travel booking was considered an important use of the Internet. Consequently an air ticketing web site search task was chosen for the study (Atterer, Wnuk, & Schmidt, 2006; Carstens & Patterson, 2005). Secondly, a single member of the community with a low level of Internet skills was observed in an air ticketing web site search task in order to signal user interface interactions that could cause difficulty in that age group. Drop-down lists appeared to present some difficulty. Small fonts also presented some difficulty. We decided to focus observations towards drop-down lists because substantial literature exists for font size.

2.1.2 The task

A task was designed to search for a return trip from an Australian capital city (Melbourne) to a regional city (Cairns) traveling on two specific dates. A particular air-ticketing site was chosen for the task primarily because it was a less popular site and consequently participants would be more likely to have a similar level of familiarity with it regardless of their knowledge of air-ticketing sites in general. Secondly it was chosen because it contained a range of form elements that operated in a similar manner to those of several major air-ticketing sites. It contained tabs for the single level overall site menu; drop-down lists for day and month/year, class of travel and passenger numbers; combo boxes for arrival and destination airports; radio buttons for type of service (flight/hotel/holiday) and trip (one-way/return/multi); and a check box to toggle flexible date search.

2.1.3 The sample

Participants were selected with the aim of obtaining a clear delineation of age group, with the control group aged between 18 and 30 (Pak R & M., 2008). The younger adults were volunteers canvassed from an Australian university with an associated Tertiary And Further Education (TAFE) or Community College level affiliation. The TAFE was
associated with a community education program from which the older adult participants were drawn. The lowest age of older adults was set to 55, partly because that was the base age in the literature (Pak R & M., 2008; Priest, et al., 2007) and partly because that was the entry age for the program. The difference of 25 years would therefore decisively delineate the two groups. University ethics compliance was simplified if participants could be drawn from university programs and if the studies took place on university grounds.

3. STAGE 1

3.1. Aims

The first stage of the study was a pilot that aimed to identify factors for further investigation in website usability for older adults by observing factors in combined multiple modalities and cognition that could reduce cognitive load.

3.2. Participants

The experimental group consisted of nine adults aged between 55 and 90. They were students in a computer class for older adults offered by a non-profit local community organisation associated with the TAFE. The class learned to use Microsoft Word, Web Browsers and email. The control group consisted of seven university students aged between 18 and 25.

3.3. Instruments

Participants were supplied with printed handouts with a minimum font size of 14 point:
   a) Ethics information
   b) Instructions for the session
   c) Pre-task questionnaire of ten items describing participant background and computing experience
   d) Task description
   e) Seven post-task questions consisting of
      • Five search result questions to record the cheapest fare and its arrival and departure times.
      • Two experience questions to record self-assessment of easy and difficult aspects of the task.

3.4. Procedure

The study took place in a computer laboratory in the university. The default browser was Internet Explorer. After confirming their understanding of the ethics and session information, participants completed the pre-task questionnaire to establish their background and computer/Internet experience and usage. They then followed the task instruction to open a browser, enter the ticketing site URL into the address bar, execute the search and answer the questions. Before the task commenced, screen capture software was
activated. On completion of the task, participants wrote down their answers to five questions about their search results and two questions about their experience of the task. Observation by the experimenters was correlated with observation of the screen capture files. The sample was considered too small for statistical analysis, which was to undertaken with the larger sample in Stage 3.

3.5. Results

3.5.1. Background information

Only one participant, an older adult, did not have a home computer and access to the Internet. All in the control group had a graduate level university qualification whereas four of the older adults had no tertiary qualification. The older adults had considerably less computing and Internet experience than the younger control group, all of whom had more than four years experience. Only two older adults had more than four years experience, with most having less than two years. They also spent considerably less time on-line. Three used the Internet less than once a day whereas all the younger adults were online at least once a day with three online constantly aided by mobile devices. The older adults mainly used the Internet for email, followed by research, whereas the younger adults were more likely to use it for social networking, followed by email, research and following the news. Several participants in each group commented that they were familiar with air ticketing web sites.

3.5.2. Task performance

When compared with the younger control group, the older adults took longer to understand the task, to execute it on the web and to check their input and the task output. Their movements were considerably slower. The older adults took far longer (between four and eight minutes, with a mean completion time of 6 minutes) than the younger group (between one minute and thirteen seconds and four minutes with a mean completion time of two minutes) to complete the task.

The most striking difference between the groups appeared in the answers to the search task questions. The entire control group correctly identified the price and flight details of the cheapest fare. However the older adults provided three different sets of flight details in their answers. Two thirds of the older adults would not have booked the cheapest fare. The screen capture files revealed that the older adults had in fact entered the air ticketing information as provided in their task instructions and had correctly interpreted the main screen which displayed their search results. They differed from the younger group in their use of the interface and its components. The air ticketing site had failed to accommodate their different behaviour. Two actions differed markedly:

• Firstly, when using combo boxes, the older adults typed the entire airport name into the input field instead of selecting from the preset drop-down list options of airport codes. This somewhat innocuous action produced an unexpected outcome – an incorrect search result of a fare of one of two higher prices, which differed depending on whether the first character of the regional flight destination airport was entered with lower case or with upper case. The capital city departure airport appeared correctly.
Secondly, only one of the older adults scrolled down below the fold line to verify the search results by viewing other fares that did not fit within the dimensions of a single screen. In contrast, all members of the control group selected from the preset drop-down list options and only one did not bother to scroll down to verify the result.

The experimental observers noted that these on-screen actions were the result of different behaviours between the groups. The younger adults, even those who did not touch type, regularly glanced at the screen while typing and entering information. They continually shifted their gaze between screen, keyboard and notes. In contrast, the older adults did not glance at the screen while typing. Instead they fixed their gaze at the keyboard while typing and then fixed their gaze at the screen for long periods of time while not typing. Before pressing the search button they spent considerable time checking what they had entered and consequently saw the ‘correct’ destination entered into the combo box.

3.6. Discussion

The length of time spent by older adults gazing at the screen could indicate that they had visual problems viewing the screen and its elements, but visual difficulties were not observed or reported. It could also have indicated that they found the task complex but they reported that the task was easy. The only difficulties reported were with a “sticky mouse” and knowing how to enter the URL into the address bar. However the observers felt that the older participants struggled to locate screen elements and form objects. Their use of drop-down lists appeared slow in comparison with the younger adults. Their inability to shift their gaze quickly between keyboard and screen could have impeded spatial and temporal contiguity, placing demands on working memory and causing them to forget the information that they were about to, or had previously, entered.

However, the most striking observation was that not one older adult appeared to know how to use combo boxes or the autocomplete option. Typing the entire field into a box appeared to be more natural for them than selecting a preset option or watching for autocomplete to present the correct option. Furthermore, only one older adult appeared to be aware of a scrolling option to view search results below the fold line.

This stage of the study was illuminating because it indicated that the major impediments to website use were not entirely perceptual-motor, but rather lack of training and experience. In summary, results of the first stage of the study were quite unexpected. They revealed that the prominent air ticketing website selected for the task contained bugs. The site had not been designed or tested for users who did not understand autocomplete and did not use preset options in combo boxes. Furthermore, typed user input had not been converted to a single case, which is a fundamental software input validation process. Consequently older adults were likely to pay considerably more than younger adults for fares and were considerably disadvantaged because of their different user interface actions.

4. STAGE 2

We were interested to discover if the faulty functionality of the air ticketing website was an anomaly or if it was in fact prevalent amongst air ticketing websites. Consequently the
second stage of the investigation compared the functionality of autocomplete in seven air ticketing websites in addition to the original site from Stage 1. It also investigated the consequences of not scrolling down by comparing the number of items and screens in the search results.

4.1. Aims

The second stage of the study aimed to compare eight well-known air ticketing websites for
a) usability and functionality of the user interface component for the flight destination airport
b) number of items and screens displayed in the search results when autocomplete was avoided in the user interface component for the flight destination airport.

4.2. Procedure

The experimental task from Study 1 was repeated with seven other websites, making a total of eight websites. The Study 2 task varied from the Study 1 task in three ways. Firstly it was completed by the experimenter and not by volunteer participants in order to systematically compare the websites by following the same steps for each site. Secondly, with each website, attempts were made to type in the destination flight airport with upper case and lower case as well as the preset-options. Thirdly, a different date applied because the time period of the first stage had expired.

4.3. Results

The results from the second stage of the study are shown in

Table 1: Comparison of attempted avoidance of preset options in online ticketing destination flight GUI components

<table>
<thead>
<tr>
<th>Site</th>
<th>User interface Component</th>
<th>Fare</th>
<th>Number of results</th>
<th>Avoidance of autocomplete</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Upper “C”</td>
<td>Lower “c”</td>
</tr>
<tr>
<td>1</td>
<td>Combo box</td>
<td>$219.00</td>
<td>$303.69</td>
<td>SAME</td>
</tr>
<tr>
<td>2</td>
<td>Text box</td>
<td>$329.00</td>
<td>$329.00</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Text box</td>
<td>$737.00</td>
<td>$737.00</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Text box</td>
<td>$928.10</td>
<td>$928.10</td>
<td>SAME</td>
</tr>
<tr>
<td>5</td>
<td>Drop-down</td>
<td>$289.00</td>
<td>$289.00</td>
<td>SAME</td>
</tr>
<tr>
<td>6</td>
<td>Text box</td>
<td>$296.00</td>
<td>$296.00</td>
<td>SAME</td>
</tr>
<tr>
<td>7</td>
<td>Text box</td>
<td>$149.00</td>
<td>$149.00</td>
<td>SAME</td>
</tr>
<tr>
<td>8</td>
<td>Drop-down</td>
<td>$449.00</td>
<td>$449.00</td>
<td>SAME</td>
</tr>
</tbody>
</table>
Although the fares differed from $149 to $928 between the websites, fares remained the same within each website, so the range of lowest fares produced by the website in Study 1 appeared to be an anomaly. However, two of the other seven websites produced a different number of results, depending on the case of the first letter typed in for the destination airport. For example, ticketing site 2 produced five results with lower case ‘c’ (Figure 1) and omitted the two results from one airline with upper case ‘C’ (Figure 2). When we repeated the test with different destinations, up to two airlines could be omitted from the search results. Although these flights had higher fares, they may have been more convenient for the user.

![Figure 1: Avoidance of autocomplete with lower case ‘c’](image1)

![Figure 2: Avoidance of autocomplete with upper case ‘C’](image2)

Of the five sites that produced consistent results, four prevented avoidance of autocomplete, two with error messages, and two by providing two drop-down lists which prevented any selection other than the preset options, one for country and one for airport within the selected country. The fifth site allowed users to type in and avoid autocomplete, but was programmed to match the airport code with the entire destination name in upper and lower case. This site produced the cheapest fare as well!

5. STAGE 3

Having confirmed within the experimental context, in possibly up to 40% of air-ticketing web sites, that older adults could be disadvantaged if they typed over autocomplete in combo and text boxes, the next stage of the study aimed to investigate the behaviour that prevented them from detecting the autocomplete options. In Stage 1 of the study, apart from spending significantly more time on task, they had been observed spending longer gazing up at the screen and down at the keyboard and their notes.

The younger adults in the first stage had a higher level of education than most of the older adults. They also had considerably more computing and Internet experience. In order to distinguish experience and educational level from factors in ageing, an effort would be made to find less educated and computer literate younger adults. This proved to be difficult.
5.1. Aims

The third stage of the study aimed to:

a) Extend the first stage of the study with a larger sample
b) Extend the first stage of the study with a control group with greater diversity of educational level and computer literacy
c) Compare older and younger adults for duration of gaze up at the screen and down at the keyboard.

5.2. Procedure

The procedures of the first stage of the study were repeated. In addition, one camcorder was placed in front of each participant to capture head movements and screen gaze duration. A second camcorder was placed to the side to capture both head and hand movements. The screen capture was synchronized with the two videos. The duration of each gaze up at the screen and each gaze down at the keyboard and notes was recorded. Results of Stage 1 and Stage 3 of the study were combined and analysed with SPSS version 19.

5.3. Participants

Eighteen older adults volunteered from another program at the same community education organisation canvassed in Stage 1. Eight younger adults volunteered from TAFE and university. Their level of education was lower than that of the Stage 1 participants but only three participants with less than four years Internet experience could be found. No participants were touch-typists. Only one of the 42 participants, an older adult, did not have home access to a computer or the Internet.

5.4. Results

5.4.1. Combined Stage 1 and Stage 3 results

It is worth noting that during the five months between Stages 2 and 3, the functioning of the combo boxes in the air ticketing web site had been corrected to process typed in cities in both alphabetic cases.

Combined results for Stages 1 and 3 were analyzed for differences in total time on task and in the number of mistakes made in reporting the search results. Significantly more time was spent on the search task \(t = 4.090, df = 40, p = .000\), two-tailed by the older adults \(m = 313.5\) seconds than by the younger adults \(m = 131.1\) seconds. Although the graduate level members of the control group in Stage 1 made no mistakes, the less educated members in Stage 3 made several mistakes in reporting the cheapest fares. Within the total sample, the older adults still made more mistakes \(m = 1.0\) than the younger adults \(m = 0.33\) when reporting their search results \(t = 2.088, df = 40, p = .022,\) one-tailed).

Participant background factors other than group were analyzed for their correlation with time on task and number of errors in reporting search results. These factors included gender, level of education, years of Internet experience, time regularly spent on the Internet. Group was broken down into its constituent age categories. Time on task was positively correlated with age category \(r = .566, n = 42, p = .000,\) two-tailed), age group \(r = .543, n = 42, p = .000,\)
two-tailed) and the number of reporting errors \((r = .313, n = 42, p = .044, \text{two-tailed})\). It was negatively correlated with time spent on the Internet \((r = -.478, n = 42, p = .002, \text{two-tailed})\) and years of Internet experience \((r = -.309, n = 42, p = .046, \text{two-tailed})\). Number of reporting errors was positively correlated with time on task and age group \((r = .313, n = 42, p = .043, \text{two-tailed})\). Internet experience was negatively correlated with age group \((r = -.582, n = 42, p < .005, \text{two-tailed})\), the finer-grained age category \((r = -.550, n = 42, p < .005, \text{two-tailed})\). It was positively correlated with time spent on the Internet \((r = .377, n = 42, p = .015, \text{two-tailed})\). There was no correlation between Internet experience and number of errors. No correlation between level of education, gender, time on task and number of errors was found.

Participant background factors were analyzed for their correlation with factors observed in the screen capture of user interface interactions. Screen capture files of two participants were corrupt leaving a sample size of 40. Use of autocomplete had a strong negative correlation with age group \((r = -.890, n = 40, p = .000, \text{two-tailed})\), age categories \((r = -.827, n = 40, p = .000, \text{two-tailed})\) and task time \((r = -.546, n = 40, p = .000, \text{two-tailed})\). It was positively correlated with Internet experience \((r = .524, n = 40, p = .001, \text{two-tailed})\) and time spent on the Internet \((r = .492, n = 40, p = .001, \text{two-tailed})\). There was no correlation between use of autocomplete, gender, level of education or number of errors. Instead of entering the supplied URL into the address bar, some older participants used a search engine to find the air ticketing website. There was no correlation between this behavior and any other factors.

Two browsers, Internet Explorer (IE) and Mozilla Firefox (MF) were used. The older adults were more likely to use IE and younger adults more likely to use MF \((r = .426, n = 40, p = .006)\). Consequently, IE users were more likely to have less Internet experience \((r = -.338, n = 40, p = .033)\) but otherwise no correlations with choice of browser were found. There was no correlation between use of the scroll bar and age group, age category, Internet experience or level of education.

5.4.2. Stage 3 results

Camcorder and screen capture files for Stage 3 only were analyzed for differences in the number of times that gaze shifted off the screen and down to the keyboard and in the duration of each gaze up and down. The number of times that gaze shifted downward off the screen was significantly greater \((t = 2.498, df = 20, p = 0.021, \text{two-tailed})\) for older adults \((m = 22.7)\) than for younger adults \((m = 14.8)\). Screen gaze duration was significantly longer \((t = 3.181, df = 20, p = 0.005, \text{two-tailed})\) for the older adults \((m = 9 \text{ seconds})\) than for the younger adults \((m = 5 \text{ seconds})\). There was no difference between the age groups in the mean duration of gazes down at the keyboard and notes.

The number of gaze shifts was positively correlated with age group \((r = .488, n = 22, p = .021, \text{two-tailed})\) and the finer-grained age categories \((r = .569, n = 22, p = .006, \text{two-tailed})\). No correlation was found with any other background or user interface factors between number of gazes or duration of gazes up or down.

5.5. Discussion

The combined results for Stage 1 and Stage 3 confirm that older adults can be significantly slower than younger adults in completing web tasks and that age is a major contributing factor. Although declines in visual, motor and cognitive skills could be associated with the
longer time spent, the results strongly indicate that less experience and less regular practice were also major contributing factors. This result is in accord with cognitive load theory which would predict that regular practice could automate many of the repeatable user interface tasks that appeared unfamiliar to the older adults yet appeared easy for many of the younger adults.

In contrast, the larger number of errors made by older adults in reporting results was correlated with no background factors other than age, indicating that either declines in cognitive skills or the cognitive demands of the task interfered with their ability to process information. Although significant, the effect was weaker than that of time on task. Possibly, a larger, less educated, control group would weaken the effect even further. Despite questionnaire responses that stated that the task was not difficult, the correlation between time on task and number of errors indicates that the longer task time was associated with difficulty in understanding and performing the task. Whereas the intrinsic load of the task may have been clearly perceived, the germaine load imposed by the user interface could have been transparent.

The argument that task complexity for older adults was related to the interface and not to the search task itself is strengthened by the Stage 3 gaze results. Duration of gazes at the keyboard and notes were similar for both younger and older adults whereas screen gaze duration was significantly longer for older adults. Screen gaze is associated to some extent with mouse and keyboard activity so the longer times could be associated with poorer motor skills. Use of some form objects, like drop-down lists, appeared slower in the older adults but not markedly slower. One older participant reported difficulty reading smaller fonts on the screen and possibly others experienced difficulty with screen layout and fonts because of poor vision. The notes were carefully formatted for older readers with large fonts, clear headings and wide spacing and should have been easy to read. If gaze duration is a measure of cognitive complexity, this result indicates that the older and younger adults had similar cognitive skills in reading and writing/typing but that the older adults had difficulty understanding and operating the user interface which was less familiar. A significantly higher number of transitions from screen to keyboard and back again for older adults indicates that information in working memory could have been lost due to insufficient spatial and temporal contiguity. The longer screen gaze durations suggest that by the time a screen element was located and its use determined, the corresponding task in the notes, or the information typed on the keyboard, could be lost from working memory and would have to be reviewed.

Use of autocomplete provided the strongest effect in the results. Only one of the 26 older adults appeared able to understand its use whereas all but one of the younger adults used it, regardless of education or Internet experience. Inability to use autocomplete was strongly correlated with longer task times. This is an important finding of this study with strong relevance for web designers and also trainers. It suggests that ways in which older adults learn to use computers differ from those of younger adults. The authors of this paper suspect that younger adults are more likely to use computers in the company of other users and consequently learn germaine interface skills through peer collaboration at stages in their learning that are conducive to automation.

In summary, a multimodal research method has produced results that confirm, refine, illuminate, and in some cases contradict, other research into factors in web usability for older adults. In addition, problems experienced with the autocomplete feature of some form elements have been uncovered.
6. CONCLUSIONS

This study investigated the interaction of different modalities and cognition in the ways older adults use websites, from the perspective of Cognitive Theory of Multimedia Learning, and its basis in Cognitive Load Theory. An iterative research methodology progressively targeted areas for further investigation while enhancing information gathered in previous stages. Complementary data collection methods enabled correlation of background survey information, test results and observation by the experimental administrators, and from screen capture and video files.

During the experimental task of a search for an airfare in an online air ticketing service web site, when compared with a younger control group, older adults spent longer on the task and made more errors when reporting their search results. The older adults, on average, had fewer years of Internet and computing experience than the younger adults. Access to a control sample of young adults with low levels of Internet experience proved problematic. However, the results were sufficiently robust to suggest that task times (but not necessarily number of errors) were strongly related to knowledge and skills within the user interface and could be reduced with regular practice and increased experience.

Firstly, time on task was strongly correlated with Internet experience and usage. Secondly, screen gaze duration can be considered a measure of attention and cognitive complexity. Older adults spent significantly longer gazing at the screen but not down at the keyboard or notes, indicating that their area of difficulty was with the interface and not necessarily with the task or decreased cognitive ability. Thirdly, the older adults did not appear to have major visual or motor skill difficulties but were completely ignorant of the use of some screen objects. Only one older adult was able to use autocomplete. More than one third did not scroll down to look at search results below the fold-line.

One consequence of this behavior was that in three out of eight air ticketing websites they would have been severely disadvantaged compared with users who used autocomplete and scroll bars. They would either have been charged higher fares or they would have seen fewer flight options. Thus, an unexpected finding from the investigation was that three out of eight prominent global air ticketing services had failed to perform basic software testing and had failed to perform usability testing with older adults or novice users. The faulty software was discovered during examination of screen capture files, without which the incorrect answers supplied by the older adults would have been attributed to ageing.

Many of the problems experienced by the older users were related to behaviours that impeded spatial and temporal contiguity. An important finding of the investigation is that older users are likely to type into combo and text boxes and to disregard preset options because they delay checking their typing until after they have completed typing the entire field. Consequently they tend to type over autocomplete options and are not able to verify their results.

Further stages in the investigation are planned, in particular with cognitive load in the use of form selection objects; reduction of spatial and temporal contiguity in mobile device touch screen keyboards; and ability to learn new web tasks, possibly within collaborative learning environments.
7. ACKNOWLEDGMENTS

The authors would like to acknowledge the support given by the Hawthorn Community Education Centre to carry out this research.

8. REFERENCES


