Web Design Guidelines for Text Presentation for Older People: Empirical Evidence from Thailand and the UK

Sorachai Kamollimsakul

Doctor of Philosophy

University of York

COMPUTER SCIENCE

September 2014

Abstract

Numerous sets of web design guidelines for making websites more accessible for older people have been suggested, but there is little empirical evidence from studies with older people upon which to base their recommendations. In addition, the different web design guidelines often provide different recommendations. Finally, most of the web design guidelines are in English and relate to the use of the Latin alphabet. Currently, there are no web design guidelines for the Thai language or for Thai older people.

The objective of this research is to investigate the recommendations from web design guidelines for Thai and UK older people, especially the recommendations related to the presentation of text for reading web pages. These are the variables investigated: line spacing, text justification, font type, font size, text colour and background colour. The recommendations were investigated with a series of empirical studies that asked both younger and older people to read web pages presented in different ways.

The first study investigated the effect of line spacing and text justification. The results of this experiment found that 1.5 or double line spacing were preferred by both younger and older people in the UK and Thailand. For the UK web readers, both left justification and left - right justification were preferred. For Thai web readers, left - right justification was preferred. As interesting issues about the task emerged in the first experiment, the second study explored the range and appropriateness of a variety of tasks for research about reading web pages. The results of the experiment indicated the use of skimming reading as an appropriate task in the further experiments.

The third study investigated the effect of font type and size on skim reading web pages. The experiment found that UK web users preferred Arial font type in comparison to Times New Roman, however Thai web users preferred a Thai conservative font type, which is closely related to serif. On font size, 14 point or larger was preferred by both the UK and Thai younger adults. For both the UK and Thai older adults, 16 point was preferred. The fourth study investigated the effect of text and background colour on skim reading web pages. Black text on white background and sepia text on off-white were preferred by all participants.

Based on results of the experiment in this programme of research, an evidence-based set of web design guidelines for the presentation of text for older people in both Thailand and the UK was developed.

Table of Contents

Chapter 1: An Introduction to the research

Chapter 2: Literature review

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

List of Tables

List of Tables (continued)

List of Tables (continued)

List of Tables (continued)

List of Tables (continued)

List of Figures

List of Figures (continued)

List of Figures (continued)

List of Figures (continued)

Acknowledgements

Author's declaration

I hereby declare that this thesis has not previously been accepted for any degree other than Doctor of Philosophy of the University of York. The research work presented in this thesis is original and my own work, otherwise indicated and acknowledged.

Parts of work contained in this thesis have been published in the following conferences:

Petrie, H., **Kamollimsakul, S**., & Power, C. (2013). Web accessibility for older adults: effects of line spacing and text justification on reading web pages. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (p. 77). ACM.

Kamollimsakul, S., Petrie, H., & Power, C. (2014). Web Accessibility for Older Readers: Effects of Font Type and Font Size on Skim Reading Webpages in Thai. In *Computers Helping People with Special Needs* (pp. 332-339). Springer International Publishing.

Kamollimsakul, S., Petrie, H., & Power, C. (2014). The Effect of Text Color and Background Color on Skim Reading Webpages in Thai. In *HCI International 2014-Posters' Extended Abstracts* (pp. 615-620). Springer International Publishing.

Chapter 1

An introduction to the research

1.1 Introduction

An important issue for many societies at the moment is the ageing of the population, with the number of older people (aged 60 years or over, see section 2.2 for a discussion of the definition of older people) rapidly increasing. The United Nations (UN, 2002) reported that the total number of older people in the world was approximately 600 million people in 2000, a threefold increase from 1950. By 2012, there were 841 million older people worldwide (UN, 2013a). The United Nations estimates that by 2050 the proportion of older people will increase to 21 per cent of the total population or more than 2 billion people (United Nations, 2002, 2013a). This will be a threefold increase from 2000. If this prediction is born out, it will be the first time in history that the proportion of the population aged 60 years and over will be larger than the proportion of young people, being those aged under 15 (UN, 2002).

In the UK, the UN (2002) estimated that the proportion of older people was 15.5 per cent of the population in 1950, rising to 20.6 per cent in 2000. By 2009 the figure had risen to 22 per cent (13.8 million) (UN, 2009). In 2013, the proportion of older people was 23.2 per cent, and the proportion of people aged 80 years and over was 4.8 per cent. It is estimated that by 2050 the proportion of older people will increase to 30.7 per cent and the proportion of people aged 80 and over will increase to 9.5 per cent (UN, 2013)

In Thailand, the UN (2002) estimated that older people was 5 per cent of the population in 1950, rising to 8.1 per cent in 2000. By 2009, the proportion had risen to 11 per cent (7.6 million). In 2013, the UN (2013) reported that the proportion of older people was 14.5 per cent, and the proportion of people aged 80 years and over was 1.9 per cent of the population. By 2050, the proportion of older people will increase to 37.5 per cent, while the proportion of people age 80 years and over will reach 10 per cent. Thailand UNFPA (2006) noted that "Thailand is ageing faster than other [countries] in South-East Asia" (p2).

The change in population demographics is also leading to an increase in the number of older web users. In the UK, 69 per cent of adults aged between 65 - 74 and 36 per cent of adults aged 75 and over have used the web (Office for National Statistics, 2014). In Thailand, the current rate of web use by older adults is very low at 2 per cent. However, the use of the web by older Thais is dramatically increasing, with a 33 per cent increase from 2008 to 2010, and a 200 per cent increase from 2010 to 2012 (National Statistical Office, 2013).

Nonetheless, older people face numerous barriers in using the web because of age-related physical, sensory, and cognitive capabilities (Holt, 2000). In addition, the lack of familiarity with the computer technologies and the web amongst the current cohorts of older people is an issue. Therefore web usability, accessibility, and user experience are important topics to empower and support older people in using websites.

According to ISO definition, ISO 9241-11 (1998), usability is defined as "the extent to which a product [or website] can be used by specified users to archive specified goals with effectiveness, efficiency, and satisfaction in specified context of use". While Petrie and Kheir (2007) explained the difference and relationship between usability and accessibility as overlapping sets which can be appeared within three types: the problems which affect only non-disabled persons are called "pure usability", the problems which affect only disabled persons are called "pure accessibility", and the problems which affect both disabled and non-disabled persons are called "universal usability". In addition, Hassenzahl and Tractinsky (2006) defined user experience as "a consequence of a user's internal state, the characteristics of the designed system, and the context within which the interaction occurs". Then the definition of "user experience" is broader than usability and accessibility as it is "all aspects of the user's experience when interacting with the product, service, environment or facility" (ISO 9241-210).

Although the Web Content Accessibility Guidelines (WCAG, 2008) are well known, they do not cover the needs of older people, only those of people with disabilities. However, a large number of web design guidelines for older people have been proposed:

- SPRY Foundation guidelines (Holt and Komlos-Weimer, 1999)
- Holt guidelines (Holt, 2000)
- Zhao guidelines (Zhao, 2001)
- AgeLight guidelines (Agelight, 2001)
- National Institute of Ageing guidelines (Hudes and Linberg, 2002)

- A checklist for the assessment Web accessibility for older users (Portuguese) (Sales and Cybis, 2003)
- AARP guidelines (Redish and Chisnell, 2004)
- SilverWeb guidelines (Kurniawan and Zaphiris, 2005; Zaphiris, Kurniawan, and Ghiawadwala, 2007)
- Webcredible guidelines (Fidgeon, 2006)

Each of these sets of guideline claims that using their guidelines will improve web accessibility and usability for older people. However, most of the web design guidelines lack evidence-based research to support their recommendations. For example, the SPRY Foundation guidelines (Holt and Komlos-Weimer, 1999) was an outcome of the conference on "Older Adults, Health Care Information, and the World Wide Web", The AgeLight guidelines (2001) were created from focus groups. The guidelines from the National Institute of Ageing (Hudes and Linberg, 2002), the AARP (Redish and Chisnell, 2004), and SilverWeb (Kurniawan and Zaphiris, 2005; Zaphiris, Kurniawan, and Ghiawadwala, 2007) were derived from reviewing and analysing the previous guidelines, in spite of the lack of empirical support for those guidelines. The Webcredible guidelines (Fidgeon, 2006) were derived from a talk-aloud usability session, while the Holt guidelines (Holt, 2000) and the Zhao guidelines (Zhao, 2001) do not provide any information about how their guidelines were developed.

Furthermore, the different web design guidelines often provide different recommendations on the same issue. For example, five sets of guidelines make recommendations about line spacing: the SPRY Foundation guidelines (Holt and Komlos-Weimer, 1999) suggest "increasing the white space between two lines of text by even a small amount (1 or 2 points)"; the Holt guidelines (2000) suggest that "older adults may have more trouble reading pages that are single-spaced rather than double-spaced. An alternative is to format paragraphs at 1½ spaces, or add a few extra points of space between lines"; the Agelight guidelines (2001) suggest that line spacing should be 2 points larger than the typeface; the SilverWeb guidelines (Kurniawan and Zaphiris, 2005; Zaphiris, Kurniawan, and Ghiawadwala, 2007) do not give specific detail, suggesting only that "there should be spacing between the lines" (p69); and the National Institute of Ageing guidelines (Hudes and Linberg, 2002) specifically suggests that line spacing must be double spaced.

In addition, most of existing web design guidelines are for English speakers reading web pages presented in the Latin alphabet. Only one set of guidelines were found for another language using the Latin alphabet, Portuguese (Sales and Cybis, 2003). No

guidelines could be found for languages which use other writing systems. Thus, it is questioned whether web design guidelines developed for the Latin alphabet be applied for other writing systems, such as that used by the Thai language? Currently, there are no web design guidelines for the Thai writing system nor for older Thai people.

Lastly, these web design guidelines were developed during the period 1999 - 2007. However, the web is a rapidly changing environment and the devices we use to access the web are rapidly changing, so new evidence is needed to provide appropriate guidelines on how to make the web accessible to older people in 2014.

1.2 Research aims and research questions

The objectives of this thesis are to investigate the recommendations from web design guidelines for older people, especially the recommendations related to text for reading web pages. In addition, the thesis investigated recommendations both for the Latin alphabet, using English speaking participants and a non-Latin alphabet, using Thai speaking participants. An empirical approach was taken, asking participants to read web pages presented with different combinations of a number of relevant variables, being:

- line spacing
- text justification
- font type
- font size
- text colour
- background colour

Both performance and preference measures were collected, being:

- Time spent per web page
- Percentage of correct answers
- Participants' rating on the visual and physical fatigue
- Users Reading Experience scores (URE)
- Participants' ratings of their overall preference

Based on the results of this programme of research, an evidence-based web design guidelines for the presentation of text for older people in English and Thai were developed.

1.3 Thesis structure

This thesis consists of seven chapters. Chapter 2 presents a review of literature related to the definitions and characteristics of older people, information about the demographics of ageing, and web design guidelines for older people. The chapter also presents previous research relevant to the recommendations on web design for older people.

Chapter 3 presents the results of the first experiment which investigated the effect of line spacing and text justification on reading web pages by younger and older people in the UK and Thailand. At the end of this chapter, recommendations on line spacing and text justification for English and Thai speaking web users are suggested.

Issues were raised in Chapter 3 about the appropriateness of the reading task used in the first experiment. Therefore Chapter 4 reports the results of the second experiment which explored the range and appropriateness of a variety of tasks for research about reading web pages. Three types of reading (scanning, skimming, and detailed reading) were investigated in the main study and searching for a link word was included in an additional round of data collection. Based on the results, skim reading was chosen to use in the two further experiments in this thesis.

Chapter 5 presents the results of the third experiment which investigated the effect of font and font size on skim reading web pages by younger and older people in the UK and Thailand. Based on the findings of this experiment, recommendations on font type and font size for English and Thai speaking web users are made.

Chapter 6 presents the results of the fourth experiment which investigated the effect of text colour and background colour on skim reading web pages by younger and older people in the UK and Thailand. Based on the findings of this experiment, recommendations on text and background colour for English and Thai speaking web users are made.

Chapter 7 presents the overall discussion of the programme of research, including the contributions of the thesis and the evidence-based web design guidelines for text presentation for English and Thai speaking people. Recommendations for the future research are also made.

Chapter 2

Literature review

2.1 Introduction

This chapter presents a review of the literature related to the characteristics of older people and web design guidelines for older people. The chapter also presents research which is relevant to the recommendations on web design for older people presented in the guidelines, with particular emphasis on research related to the presentation of text on web pages.

The structure of this chapter is as follows: Section 2.2 presents definitions of older people, including those proposed by international organizations and in the research literature. Section 2.3 presents information about the demographics of ageing. Section 2.4 presents the characteristics of older people and follows with details about cognitive and physical capabilities in older people in Section 2.5. Section 2.6 presents the various web design guidelines for older people and Section 2.7 presents specific details about web design guidelines related to changing visual capabilities in old age. Section 2.8 presents previous research on recommendations in the web design guidelines related to the presentation of text on web pages, specifically line spacing, text justification, font type, font size, and text and background colour.

2.2 Definitions of older people

There are different definitions of who "older people" are from different organizations and research groups. This section reviews the definitions of older people used by the World Health Organization (WHO), the United Nations (UN), and a range of research studies. It also provides information on the demographics of the older population in both Thailand and the UK.

The term "older people" has a number of different definitions. Many organizations make different suggestions for the age at which people shift from being "adults" to "older adults". According to the World Health Organization (WHO, 2000), the United Nations

(UN) does not specify a minimum chronological age for older people. However, Ferreira and Kowal (2006) noted that a minimum age of 60 years in used in most UN publications. The WHO (2000) itself suggests that the minimum age of 65 years is selected by most developed countries, as it is the typical age at which men retire (Arch, 2008). However, women have typically retired earlier in the UK and with the ageing population, retirement ages across the developed world are now changing. The minimum age used by the WHO is in fact varied: as Kowal and Peachey (2001) note, in the 2000 Harare Minimum Data Set Workshop a minimum age of 60 years was used, but during the 2001 Dar es Salaam Minimum Data Set Meeting this was changed to 50 years old because this age better represented a realistic definition of older people in the developing countries. Other organizations also use 50 years as the minimum age, even in the developed world, including the American Association of Retired Persons (AARP, 2014). In summary, each organization defines older people as a person over a different age. 50, 60, and 65 years are all used by international and national organisations.

Table 2.1: Age ranges (in years) for young, middle aged and older people in the research literature (Nichols et al 2001)

	Human Factors Journal 1998 – 2000	**Psychology and Aging Journal 1995 – 1999**
Young people	19.1 – 34.6 years	18.9 – 30.1
Middle aged people	39.7 – 58.7	40.9 – 57.2
Older people	57.5 – 76.1	57.3 - 62.1

Older people are also variously defined in the research literature as presented in Table 2.1. Nichols et al (2001) found 131 articles in the Human Factors Journal published between 1998 and 2000 that participants classified as "young people" had a mean age range of between 19.1 and 34.6 years, "middle aged people" had a mean age range of between 39.7 and 58.7 years, and "older people" had a mean age range of between 57.5 and 76.1 years. There is no place for people who are aged 34.7 and 39.6 years, and there is also some overlap between ranges used for middle aged and older people. On the other hand, Nichols et al (2001) found 202 articles in the Psychology and Aging Journal published between 1995 and 1999 that participants classified as "young people" as those who had a mean age range of between 18.9 and 30.1 years, "middle

aged" had a mean age range of between 40.9 and 57.2 years, and "older people" had a mean age range of between 62.2 and 82.3 years. This set of groupings also has the gap between each group, with no place for those aged 30.2 to 40.8 years and 57.3 to 62.1 years old. Bailey (2004) suggested that researchers in the field of ageing should adopt a consistent age classification. Bailey proposed that the most appropriate categories for age groups are: "young" 18 to 39 years; "middle-aged" 40 to 59 years; "older" 60 to 74 years. In addition, Bailey suggested the "old – old" for people aged 75 years and over.

Other researchers in the field of ageing also break the older group of people into finer grained categories. For example, Garfein and Herzog (1995) and Chi and Chou (2002) divided older people into three groups: "young-old" - 60 to 69 years; "old-old" - 70 to 79 years; and "oldest-old", 80 years and over. Spirduso, Francis, and MacRae (2005) separated older adults to four categories: "young-old" - 65 to 74 years, "old" - 75 to 84 years, "old-old" - 85 to 99 years, and the "oldest-old", 100 years and over.

After reviewing the research about how older people use Information and Communication Technologies (ICTs) and the web that was published between 2000 and 2004, Redish and Chrisnell (2004), concluded that researchers used many different definitions of older people. The lowest minimum age of older people in research was 50 years, while the highest minimum age was 70 years.

Table 2.2, below, shows the minimum age of participants in 45 ICT related studies of older people conducted between 1985 and 2009 which the current author sampled at random. The analysis presented in the table agrees with that given by Redish and Chrisnell (2004) that the definition of older people varies widely between studies. For my analysis, the lowest minimum age for older people was 44 years, while the highest minimum age was 70 years. Researchers most frequently identified "older people" as people were aged over 65 years (10 papers), over 60 years (7 papers), over 55 years (5 papers), or over 50 years (5 papers).

Table 2.2 The lowest age of participants in 45 ICT studies about older people reviewed by the current author

The minimum age (in years) of older participants	Number of research papers	References
44	1	Gao et al. (2007)
50	5	Maguire and Pearce (2001) Wright and Belt (2001) Chisnell, Lee and Redish (2004) Moore and Matthews (2004) Ryu et al. (2009)
51	1	Turns and Wagner (2004)
52	1	Wolters et al (2009)
53	2	Aula (2005) Kim et al. (2005)
54	1	Jacko et al. (2002)
55	5	Lines and Elliman (2007) Chawick-Dias, McNutty and Tullis (2003) Chadwick-Dias, McNutty and Tullis (2004) Wang et al. (2007) Moffatt and McGrenere (2007)[1]
56	1	Aula and Kaki (2005)
57	2	Moscicki et al. (1985) Lin (2003)
59	1	Marquie et al. (2002)
59	1	Dror et al. (1998)
60	7	Czaja and Sharit (1998) Smith et al. (1999) Morrell, Mayhorn, and Bennett (2000) [1] Czaja et al. (2001) Hawthorn (2003) Pfeil et al. (2009) Struve and Wandke (2009)

The minimum age (in years) of older participants (continued)	Number of research papers (continued)	References (continued)
61	1	Wright, Belt and John (2004)
62	2	Bernard, Lio and Mills (2001) Fukuda and Bubb (2003)
64	1	Groff et al. (1999)
65	10	Huey et al. (1996) Chaparro et al. (1999) Smith et al. (1999) Coyne and Nielsen (2002) Nielsen (2002) Kantner and Rosenbaum (2003) Keates et al (2004) Fidgeon (2006) Sayago et al. (2009) Struve and Wandke (2009)
66	1	Moffatt and McGrenere (2009)
70	3	Koyani et al. (2002) Keates and Trewin (2005) Moffatt and McGrenere (2007) [1]
75	1	Morrell, Mayhorn, and Bennett (2000) [1]

1. There were two studies with different minimum age of participants in this paper.

However, some research which has investigated older people in more than one country has used different minimum ages for older participants in each country. For example, Malik (2011) conducted a study about older people, mobile technology, and culture. The participants in the research were older people in the UK and Malaysia. An interesting point in this study was the method Malik used to calculate the appropriate minimum age for the older participants in the two countries. Malik noted that each country has different retirement age and life expectancy. These differences mean that people spend different proportions of their lifes in retirement. For example, 65 years is currently the typical retirement age in the UK for men and the average life expectancy of UK men is 80 years. This means that UK men on average spend 19% of their lives

in retirement. Whereas, for Malaysian men the average retirement age is 58 and their life expectancy is 78 years. So they spend 26% of their lives in retirement. Malik set the average proportion of life in retirement for UK men as a benchmark and then calculated the age at which Malaysia people should retire from work in order to spend the same proportion of life in retirement as UK men.

Table 2.3: Method for calculating minimum age for in different countries used by Malik (2011)

	Retirement Age (RA)	Life Expectancy	Retirement Proportion	Adjusted Retirement Age (ARA)
UK (men)	65	80	19%	65
Malaysia (men)	58	78	26%	59.5

Malik's calculations are shown in Table 2.3., The minimum age of older participants in the UK is 65 years and the minimum age of older participants in Malaysia is 59.5 years. However, this method used retirement age of UK men only as a benchmark, but there are different retirement ages for men and women in the UK. In addition, retirement ages in each country are currently changing, as noted above, with the UK and many countries now delaying retirement age for both men and women. However, this general method is useful for comparing the minimum age for older people in different countries.

In summary, this section has shown that there is evidence that the definition of older people varies considerably between and even within international and national organizations. The suggestions about age minima from different researchers also vary, and there is no final agreement. However, after reviewing the research about older people and ICT, it was found that the most commonly used minimum age for older people is 65 years, while 60, 55, and 50 years are also commonly used. In addition, some research which had participants in more than one country defined the minimum age of older participants at different ages for each country, supported by appropriate calculations.

For this research programme, the definition of younger people proposed by Bailey (2004) of 18 to 39 years old, was adopted for use both in Thailand and the UK as it was suggested from reviewing many research papers. Bailey's categorization also provided a full set of age ranges, with no gap or overlap between each age. While the minimum age for older people in the UK adopted was 65 years as it is typical age of research

about older people and ICTs. In addition, 65 years old is also one of breakpoints between categories of older people used by many researchers including Chi and Chou (2002), Garfein and Herzog (1995), and Spirduso, Francis, and MacRae (2005). However, the minimum age for older participants in the research in Thailand was 58 years, because at this age older participants in Thailand have the same average proportion of life in retirement as older participants in the UK (see section 2.3 Demographics of Ageing).

2.3 Demographics of Ageing

An important issue for many societies at the moment is the ageing of the population, with the number of older people rapidly increasing. The United Nations (UN, 2002) reported that the total number of people over the age of 60 in the world was approximately 600 million people in 2000, a threefold increase from 1950. In 2012, there were 841 million older people worldwide and it is estimated that the number of older people will increase to 2 billion by 2050 (UN, 2002, 2013a). Again, this will be a threefold increase from 2000. Figure 2.1 shows the world population pyramids for the years 1950, 2000, and 2050.
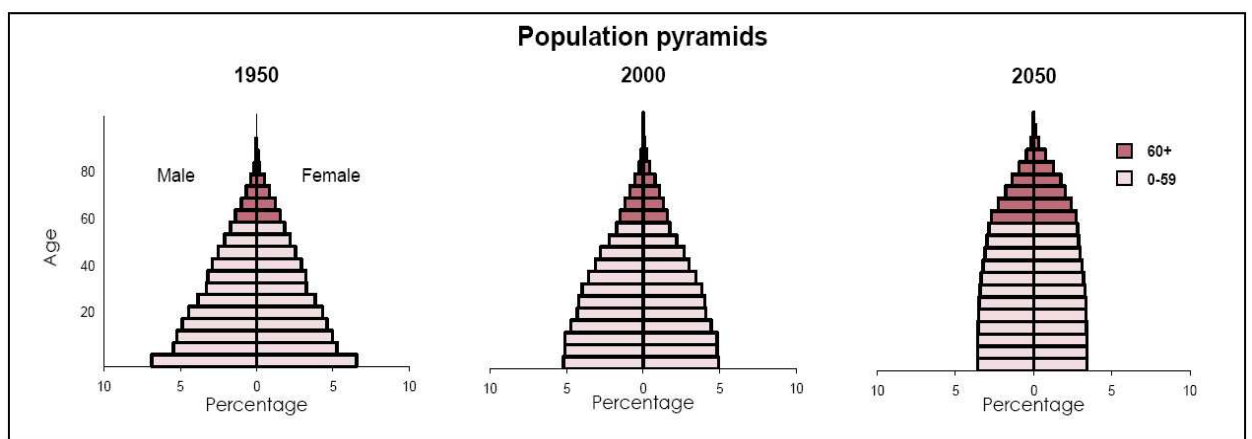
Figure 2.1: World population pyramids for 1950, 2000 and 2050

(Source: United Nations, 2002)

Not only is the number of older people increasing, but perhaps more importantly, the proportion of older people in the population is also increasing. The proportion of older people was only 8 per cent in 1950, rising to 10 per cent in 2000, 12 per cent in 2013 and it is forecast to be 21 per cent by the year 2050 (UN, 2002, 2013), as shown in

Figure 2.2. This situation is unprecedented in the history of humankind. In parallel to the increase of the proportion of older people in the population is a corresponding decrease in the proportion of young people. By 2050, it is estimated that it will be the first time in history the proportion of people aged over 60 years will be larger than the proportion of young people aged under 15 years (UN, 2002).
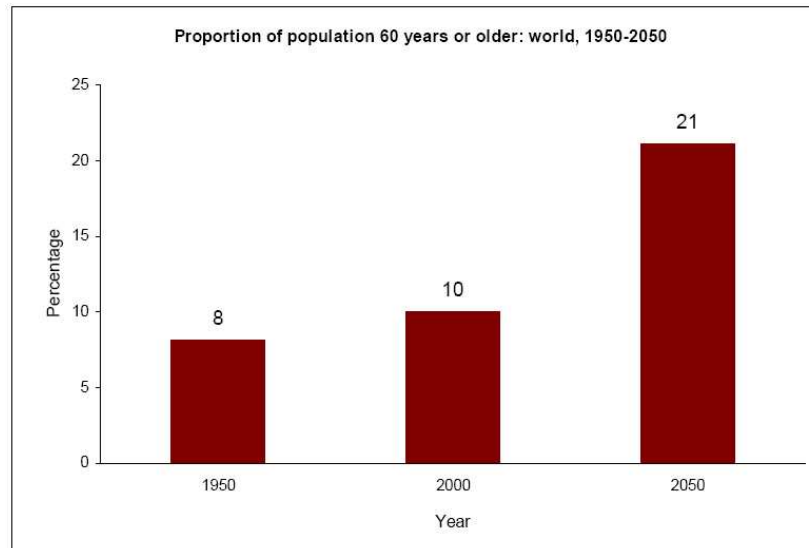
Figure 2.2: Proportion of world population 60 years or older for 1950, 2000 and 2050 (Source: United Nations, 2002)

Figure 2.3 shows the population pyramids of less and more developed regions. The less developed regions comprise all regions of Asia (excluding Japan), Africa, Latin America and the Caribbean, and Oceania (excluding Australia and New Zealand). The more developed regions comprise all other regions of the world and the three countries excluded from the less developed regions. In recent statistics, UN (2013b) reported that the number of older people in less developed countries is growing faster than in developed countries. In less developed countries, there were 554 million older people in 2013, five times higher than in 1950. It is predicted that it will increase three fold by 2050, to reach 1.6 billion. The growth in the number of older people in developed countries is slower than in less developed countries, however, it is still a very important change. The number of older people in developed countries was 94 million in 1950, increasing threefold to 287 million people in 2013, and it is predicted that it will reach 417 million people by 2050.
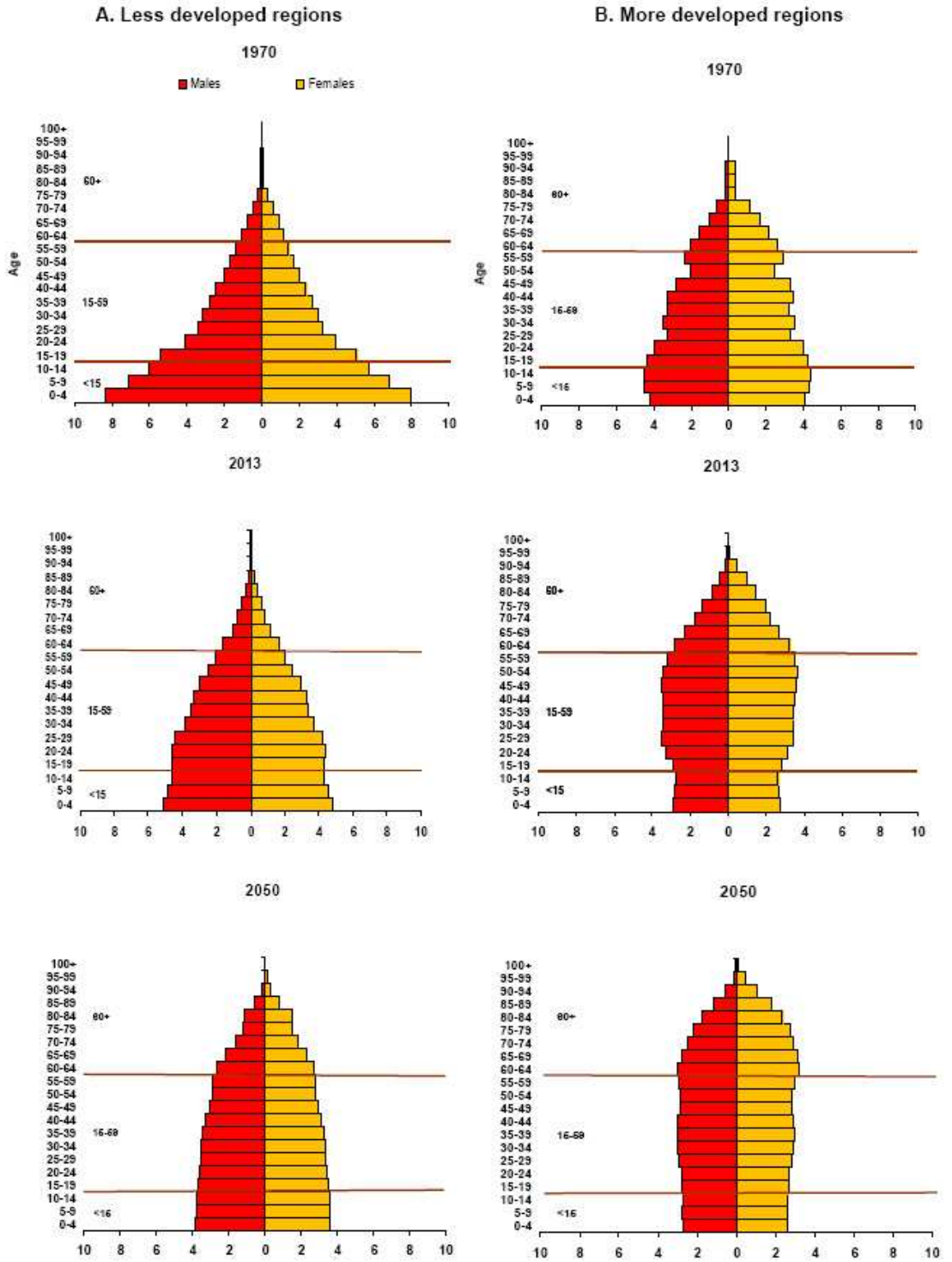
Figure 2.3: Population pyramids of less and more developed regions

(Source: United Nations, 2013b)

In the UK, the UN (2002) estimated that the proportion of older people was 15.5 per cent of the population in 1950, rising to 20.6 per cent in 2000, and that it will increase to 34 per cent by 2050. Figure 2.4 illustrates the relevant population pyramids of the UK. In 2013, the proportion of people aged over 60 was 23.2 per cent, and aged 80 years and over was 4.8 per cent (UN, 2013a).
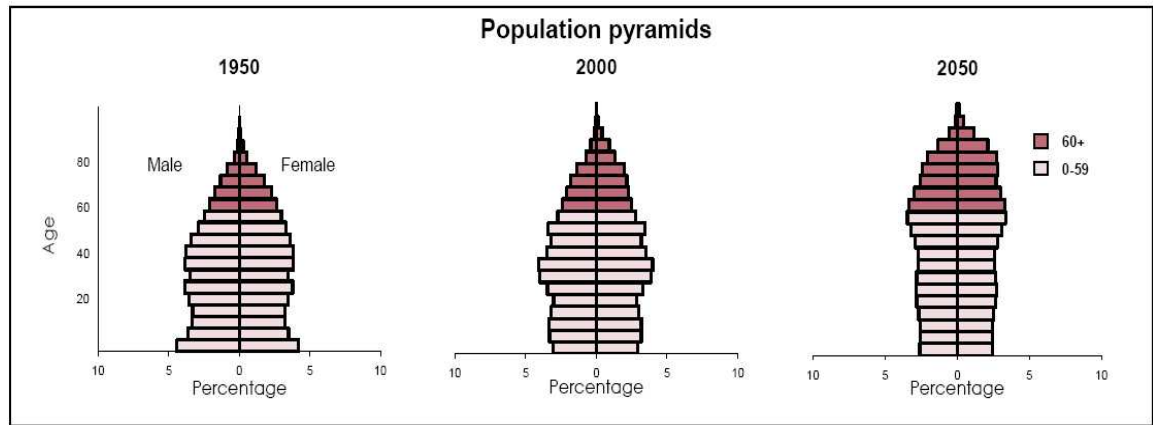
Figure 2.4: Population pyramids of United Kingdom of Great Britain and Northern Ireland (Source: United Nations, 2002)

In Thailand, the UN (2002) estimated that people aged 60 years and over was 5 per cent of the population in 1950, rising to 8.1 per cent in 2000, and 27.1 per cent in 2050. Figure 2.5 illustrates the relevant population pyramids of Thailand. In 2013, the UN (2013) reported that people aged 60 years and over was 14.5 per cent, and people aged 80 years and over was 1.9 per cent of the population. UNFPA Thailand (2006) noted that "Thailand is ageing faster than other [countries] in South-East Asia" (p2). The population pyramids of Thailand are greatly changing.
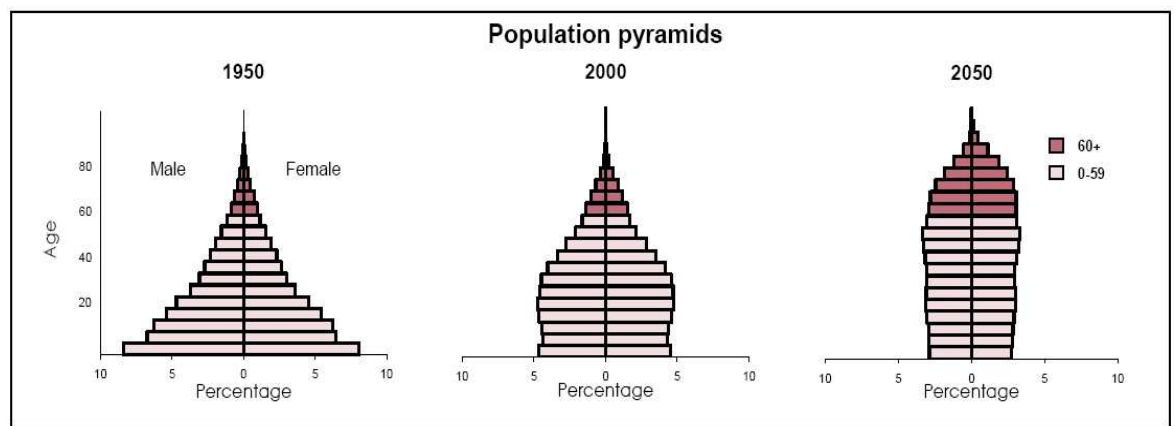
Figure 2.5: Population pyramids of Thailand (Source: United Nations, 2002)

Another important indicator about the ageing population is the population dependency ratio. The population dependency ratio is the proportion of younger people (up to 14 years old) and older people (65 years old and over) to the number of people of working age (15 to 64 years old) (Barcelona Field Studies Centre, 2011). This is important as it shows the proportion of the population generating income and wealth and caring for both older people and children in relation to the proportion needing care.

Table 2.4: Old age dependency ratio for the world,
United Kingdom and Thailand

(Source: United Nations, 2002)

Country	Old age dependency ratio		
	1950	**2000**	**2050**
World	8.6	10.9	24.7
United Kingdom	16	24.1	47.3
Thailand	5.9	7.7	34.1

The population dependency ratio can be divided into the youth dependency ratio and the old age dependency ratio, reflecting the ratio of young people to people of working age and older people to people of working age respectively. As summarised in Table 2.4, globally the old age dependency ratio was 8.6 in 1950, rising to 10.9 in 2000, with a further gradual rise to 24.7 predicted by 2050. The old age dependency ratio in the UK was 16.0 in 1950, 24.1 in 2000, and is predicted to double to 47.3 by 2050. In Thailand the old age dependency ratio was 5.9 in 1950, rising to 7.7 in 2000, and predicted significantly increase to 34.1 by 2050 (UN, 2002).

Table 2.5: Life expectancy for the world, United Kingdom, and Thailand

(Sources: United Nations, 2002, 2013a)

Country	Life expectancy					
	1950 - 1955	1975 - 1980	2000 - 2005	2010 - 2015*	2025 - 2030	2045 - 2050
World	46.5	59.8	66.0	70.0	72.4	76
United Kingdom	69.2	72.8	78.2	80.4	81.4	83.0
Thailand	52.0	61.4	70.8	74.3	76.8	79.1

There are two important indicators which relate to the ageing population; life expectancy and healthy life expectancy. Life expectancy is defined as the average age that a person may expect to live (WHO, 2006). Healthy life expectancy is defined as the average age that a person may expect to live with full health (WHO, 2006).

Table 2.5 shows that the average life expectancy of the world population increased by 20 years during the period from 1950 to 2000, from 46.5 years to 66 years, and is expected to increase by a further 10 years between 2000 and 2050, from 66 years to 76 years. In Thailand, the life expectancy has increased at nearly the same rate as the overall world figures. It increased approximately 20 years between 1950 to 2000, and will increase by approximately 10 years between 2000 to 2050. In the UK, life expectancy increased only 10 years from 1950 to 2000 and a further five years from 2000 to 2050. Thus, the gap in life expectancy in developed countries such as the UK and developing countries such as Thailand will decrease in the near future.

Table 2.6: Healthy Life expectancy in the United Kingdom and Thailand

(Source: Global AgeWatch, 2013)

Country	Healthy life expectancy
United Kingdom	71
Thailand	60

Table 2.6 shows the healthy life expectancy for the UK and Thailand. It shows that people in developed countries such as the UK can live with full health longer than

people in developing country such as Thailand. People in the UK and Thailand, were expected to live with full health until they are aged 71 and 60 years, respectively.

For conducting research with older people in two different countries, the UK and Thailand, the current researcher was concerned about how to select an appropriate minimum age for older participants. As shown in Table 2.4 and Table 2.5, the UK has higher ages for both life expectancy and healthy life expectancy than Thailand. These mean that older people in the UK are expected to live longer than older people in Thailand. Not only do they live longer, but older people in the UK also have full health for longer. If the same age of older participants in the two countries were to be used in this research programme, such as 65 years, it would mean that it would be comparing results between older people who were have their full health (i.e. those in the UK) with older people who are no longer in full health (i.e. those in Thailand).

This literature review suggested that healthy life expectancy should be considered as a factor for deciding the minimum age of older participants in each country. To calculate appropriate minimum ages for older participants in the current programme of research, the minimum age of participants in the first country (the UK) was set and then the minimum age of participants in the second country (Thailand) was calculated as that yielding the same proportion healthy life expectancy. The minimum age of older participants in the UK, was set at 65, as this has been the typical retirement age in western countries for some time (although it is now changing) and this is the most frequently used minimum age for older people in research relating to ICTs (see section 2.2). The appropriate minimum age for the second country, Thailand, was calculated the following formula:

$$\text{Appropriate minimum age for the second country} = HLE1 - \left(\left(\frac{HLE1 - RA1}{HLE1}\right) * HLE2\right)$$

RA1: Participants' minimum age for first country (the UK)

HLE1: Healthy Life Expectancy in the first country

HLE2: Healthy Life Expectancy in the second country

Table 2.7: Figures for calculating minimum age for the UK and Thailand older participants for this research programme

	Healthy Life Expectancy (HLE)	Appropriate minimum ages for older participants
United Kingdom	71	65 (set)
Thailand	60	54.9 (calculated)

Table 2.7 shows the figures for calculating minimum age for older participants for the UK and Thailand. From the calculation, participants in Thailand age at 55 years had the same proportion of their remaining life as healthy as participants in the UK at 65 years. This age, 55 years, is also one of the frequently used minimum age for older people in the research literature. Thus, in this programme of research, the older participants in the UK were people aged 65 years and over while older people in Thailand were people aged 55 years old and over. With these two different minimum ages for older participants, the results from the two countries can be compared more equitably.

In summary, the proportion of older people who aged 60 years old or over is dramatically increasing and will continue to increase in the future. There will be 2 billion older people in the world by 2050. In addition, it will be the first time in human history that the proportion of older people aged 60 years old or over will greater than the proportion of younger people aged less than 15 years old. Moreover, the old age dependency ratio has already risen dramatically since 1950 and will continue to increase in many countries. All these factors mean that older people are more important to study than the past. In this section, ages of older participants for the empirical studies in this programme of research were set. The older participants in the UK are aged 65 years and over while the older participants in Thailand are aged 55 years and over.

2.4 Characteristics of older people

After reviewing much research on older people, Redish and Chrisnell (2004) warned that stereotyping older people as a single group by age would mean that researchers miss important design features that would benefit older people because older people are so varied in their characteristics.

Gregor, Newell, and Zajicek (2002) divided older people into three groups depending on their physical and cognitive abilities. These are: fit older people, frail older people, and disabled people who grow older. Fit older people are older people who do not have any disabilities and believe themselves not to have any disabilities. However, older people in this group are usually weaker than when they were younger. Frail older people are older people who have one or more disabilities, limitations, or decline in their sensory abilities. Disabled people who grow older are older people who have a long term disability or disabilities which may be further affected by ageing.

Apart from chronological age, some researchers are concerned about other characteristics of older people and noted some important characteristics of older people in their research, such as computer expertise and experience and web expertise and experience (Redish and Chrisnell, 2004; Gregor, Newell, and Zajicek, 2002),

Redish and Chrisnell (2004) introduced a new approach with four dimensions for categorising older people. There were age, ability, aptitude, and attitude. Age is defined by them as both chronological and experiential, including maturity level, which they defined as life events and experiences. Ability they defined as levels of physical and cognitive limitations. Aptitude they defined as levels of expertise with computers and the Web. Attitude they defined as confidence levels and emotional state of mind. In order to understand about these four attributes, Redish and Chrisnell asked web designers to read personas about older people and then decide where to put each persona in the levels for each attribute. Redish and Chrisnell argued that these four attributes are useful to judge the levels of support and training particular older people might need and the levels of complex features in computing systems that they can deal with.

These four attributes are interesting in terms of a novel method to classify older people. However, there have some difficulties in using them, for example, how many levels of differentiation are appropriate for each attribute? In addition, chronological age is easier to measure, but any measure of ability, aptitude or attitude is much more complex. Moreover, ability, as defined by Redish and Chrisnell, included both physical and cognitive ability, which are different concepts and both complicated to measure. Finally aptitude and attitude are also challenging multi-dimensional attributes.

There have been numerous scales to measure computer anxiety and attitudes (Davis,1989,1993; Heinssen, Glass and Knight, 1987; Loyd and Gressard, 1984; Nickell and Pinto, 1986). However, those measurements tend to measure specifically

anxiety and attitudes toward computers and now quite dated. Burn (2003) developed a scale to measure attitudes toward technologies in a general standardized manner and used factor analysis to extract the different factors underlying the attitudes. Burn's scale comprises 18 questions grouped into three factors: Confidence, Performance, and Fashion. The Confidence Factor contains nine questions that relate to how easy it is to learn and to remember to use the technology, and one's own confidence and difficulty in using the technology. The Performance Factor contains six questions that relate to the performance of the technology: its efficiency, effectiveness, and reliability, and its perceived value for money. The Fashion Factor contains three questions about the effect of using the technology in terms of being a fashionable person, whether the technology is positive to the person's image, and usage by peers. Thus, the attitudes of older people toward technologies such as the web can be measured in a more detailed manner than simply using a single question as suggested by Redish and Chrisnell (2004).

2.5 Physical and cognitive changes in older people

The WHO (2012) notes that as people are living longer in our ageing society, more people are likely to have to deal with disabilities when they get older. Goodman-Deane, Keith, and Whitney (2009) also noted that "age on its own is not a disability, but older people are more likely to experience disabilities of various kinds" (p1). Moreover, Monk (2009) supported the notion that aging is a cause of changes in human abilities. The reason is older people experience losses in their physical and mental abilities due to the normal aging process (Blaschke, Freddolino, and Mullen, 2009). Arch (2008) concluded that the limitations according to age, which have an impact on access to technology use are changes in vision, hearing, motor and cognitive abilities.

The next sections will review the changes that people experience as they age in each of these areas and the consequences of these changes for web use.

2.5.1   Changes in vision due to ageing

The aspects of change in visual processing that cause older people problems are near object focusing ability, changes in colour perception and sensitivity, changes in contrast sensitivity, and reduction in visual field (Agelight, 2001; Salvi, Akhtar and Currie, 2006). The Royal National Institute for Blind People (RNIB) (2014a) notes that diseases such

age-related macular degeneration (AMD), cataract, diabetic retinopathy, glaucoma are leading causes of sight loss.

The RNIB (2014b) reported statistics about people with sight loss in the UK, noting that there are nearly two million people facing with sight loss, about one in 30 persons. People with all ages are affected by sight loss, but older people are much more frequently affected, with 80 percent of those with vision loss being over the age of 60 years. The percentage increases as people get older, with 20 per cent of people aged 75 years and older having substantial vision loss, and 50 per cent of people aged 90 years and over having substantial vision loss. It is predicted that the number of people with sight loss in the UK will increase to more than 2.25 million people by 2020. The population with sight loss will continue to increase to nearly 4 million by 2050.

There was no specific research showed the results how changing in vision due to aging affected using the web. However, changing in vision due to aging seem to have great impact on using the web as the results h from the Disability Rights Commission (DRC) research. Although, the participants in DRC research were not older people, they were disabled persons. However, the disabilities has related to some other age related declines in older people. In that research, the DRC (2004) conducted a series of studies about web accessibility. The participants in the main user testing study were 51 people with a variety of disabilities, including blindness, partially sightedness, dyslexia, profound deafness, and physical impairment. Each participant was asked to evaluate 10 web sites and undertake two tasks on each website. The results shown that the average task completion rate of participants with all types of disabilities was 76 per cent. However, blind participants were the least successful in completing tasks with only a 53 per cent success rate, participants with partially sighted were more successful at 76 per cent, while the other groups of disabilities had success rates over 80 per cent. In addition, participants were asked to rate how easy it was to complete the tasks. The blind participants and participants with partially sight rated the tasks more difficult than others. It might be implied that older people with vision loss will face more difficulties in accessing the web than people with other age related declines.

Dickinson et al. (2005) concluded that older people's vision loss is the reason for many of their difficulties in the use of technology such as reading text labels, buttons, and problems of screen contrast. AgeLight (2001) also argued that of the changes due to ageing, the greatest impact is those in vision.

2.5.2 Changes in hearing due to ageing

Hearing loss is one of the physical changes that affect people as they age. It causes many difficulties in communication. According to the Royal National Institute for Deaf People (RNID, 2011), there are four different levels of hearing loss: mild hearing loss, moderate hearing loss, severe hearing loss, and profound Schwartz (2012) explains that there are tiny hair cells within inner ears, which pick up sound wave and change to nerve signals then the brain acknowledges that sound. Age-related hearing loss typically occurs when these hair cells die or are damaged. Other reasons for age-related are that the three tiny bones inside the ear can no longer conduct sound properly or that other structures in the ears are damaged (Vorvick, 2012).

Older people with hearing loss usually experience difficulty in hearing other people with higher-pitched voices, problems hearing in noisy environments, more frustration because of not being to hear than when they were younger, and ringing sounds in the ears (Schwartz, 2012, Vorvick, 2012).

The RNID (2011) reported that more than 10 million people in the UK had some form of hearing loss. Most of these people, approximately 6.4 million, are older people aged 65 years and over, in comparison to only 3.7 million people aged 16 to 64 years. In addition, the statistics showed that approximately 40 per cent of people aged 50 years and over and 70 per cent of older people aged 70 years and over have some degree of hearing loss (RNID, 2011). In addition, Mehta (2014) noted that approximately 90 per cent of people aged 80 years and over have some form of hearing loss. [I would omit this sentence, you have established this very clearly already and this does not look like an academic reference, and there is not reference: Hearing Link (RNID, 2011) noted that increasing of hearing loss is significantly related to increasing age. It is predicted that the number of people with hearing loss in the UK will increase to approximately 14.5 million people by 2031.

In terms of the effect of hearing loss has on using the web for older people, it is currently not considered a particular barrier due to the largely visual nature of the web (DRC, 2004; Hanson, 2001). However, as the web becomes more multimedia with videos and audio information, this could change.

2.5.3   Changes in motor skills due to ageing

Changes in motor skills due to ageing include slower response times, loss of flexibility, disturbances in coordination, decreasing ability to balance during continuous movements, and less accurate and more variable movement (Czaja and Moen, 2004; Seidler et al., 2010). The main diseases of ageing that lead to changes in motor skills are arthritis and Parkinson's Disease (Arch, 2008; Ilyas, 2012).

Problems related to computer use for older people with problems with motors skills include controlling a mouse and other input devices. Bohan and Scarlett (2003) conducted a study about the effect of expanding targets on the object selection performance of older adults. 8 younger and 8 older participants were asked to complete target acquisition tasks in five conditions: small static, large static, 10% expansion, 50% expansion, and 90% expansion. The results showed that the older participants spent significantly longer time to acquire the target than younger participants across all conditions. In addition, there was a significant main effect of target condition that suggested that expanding targets were a technique which can improve older persons' performance on target selection.

Keates and Trewin (2005) conducted a study about cursor positioning using a mouse. There were 31 participants, including young adults (20 to 30 years), middle-aged adults (35 to 65 years), older adults (70 years and older), and adults with Parkinson's disease (48 to 63 years). They found that older people and adults with Parkinson's took longer time to complete the experimental task. In addition, they found that the average number of pauses per movement increased with increasing age, while the average number of pauses per movement of adults with Parkinson's was between the averages for adults and older adults.

In addition, Dickinson et al. (2005) also mentioned that changes in motor skills in older people have substantial effects on mouse use and mouse control, particularly in locating small targets.

2.5.4   Changes in cognition due to ageing

Age-related changes in cognition include diminished capacity of working memory; reduction in the ability to learn and remember new information; reduced cognitive processing speed; decline in spatial and visual information processing; decreased ability in dividing attention between two or more tasks, and a greater chance of

experiencing interference in long-term memory (Kurniawan et al., 2006; Czaja and Moen, 2004). Arch (2008) noted that dementia, including Alzheimer's disease, and Mild Cognitive Impairment are common causes of cognitive impairment in older people.

In 2012, there were 800,000 people with dementia in the UK (Alzheimer's Society, 2012a). About five to 20 per cent of older people have some symptoms of mild cognitive impairment and about 10 to 15 per cent of people with mild cognitive impairment went on to develop dementia (Alzheimer's Society, 2012b).

Meyer et al. (1997) conducted a study about age differences in web navigation. There were 20 participants: 13 older and 7 younger adults. The author did not provide participants' specific details. The participants were asked to search a complex web site to find a specific piece of information. The results showed that older participants took significantly more steps for finding the answer than younger participants.. Older participants also tended to return to the homepage more often to complete the task. Finally, older participants returned to web pages which they had visited more often than younger participants. The researchers inferred that the older participants could not remember which web pages they had visited and also could not remember the information on those web pages as well as younger participants.

Thus, not only changes in physical abilities have an effect on older people using the web, but changes in cognition due to aging also has an important effect.

Therefore web accessibility and usability is an important topic for empowering and supporting older people to be able to use websites. Many researchers and organisations suggest the guidelines for making website easier to use for older people. The section 2.6, below, provides more specific details on web design guidelines for older people.

2.6 Web design guidelines for older people

As discussed in Section 2.5, above, people experience changes in their physical and cognitive abilities as they age. Web design guidelines are one important solution to help web developers create websites that will be easy for older people to use to overcome these changes. Web accessibility guidelines can be divided to two types: general web accessibility guidelines and web accessibility guidelines for older users.

The main set of general web accessibility guidelines, the Web Content Accessibility Guidelines, is currently in its second version (WCAG 2.0). WCAG was developed by

the Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C). WCAG 1.0 was published in 1999 in order to explain to web content developers how to create web sites usable by people with disabilities. This original set of guidelines has 14 guidelines which break down into 65 checkpoints. Each checkpoint has a priority level: Priority 1 means that this checkpoint must be satisfied by web content, otherwise according to WAI, one or more groups of disabled users will find it impossible to access information in the document. Satisfying this checkpoint is a basic requirement for some groups to be able to use Web documents. Priority 2 means that web content should be satisfy this checkpoint. Otherwise, according to WAI, one or more groups of disabled users will find it difficult to access information in the document. Satisfying this checkpoint will remove significant barriers to accessing Web documents. Finally Priority 3 means this checkpoint may be satisfied by web content. Otherwise, one or more groups of disabled users will find it somewhat difficult to access information in the document. Satisfying this checkpoint will improve access to Web documents. There are three levels of conformance: level A (Priority 1 checkpoints are satisfied), AA (Priority 1 and 2 checkpoints are satisfied), and AAA (Priority 1, 2, and 3 checkpoints are satisfied), (Chisholm et al., 1999).

WCAG 2.0 was published in 2008. This revision of the web accessibility guidelines updated and expanded WCAG 1.0 and attempted to be less technology specific and more "future proof". In WCAG 2.0, there are 4 four principles: web content should be perceivable, operable, understandable, and robust. There are 12 guidelines nested under these 4 four principles. As with WCAG 1.0, each guideline has a success criterion: level A (lowest), AA, and AAA (highest). WCAG 2.0 still aims to increase accessibility of the web to people with different type of disabilities: blindness, deafness and hearing loss, learning disabilities, cognitive limitations, limited movement, speech disabilities, photosensitivity and combinations of disabilities (Caldwell et al., 2008).

However, this programme of research is particularly interested in web accessibility guidelines for older users. A number of sets of web accessibility guidelines for older people have been found, as follows:

SPRY Foundation guidelines (1999)

Holt guidelines (2000)

Zhao guidelines (2001)

AgeLight guidelines (2001)

National Institute of Aging guidelines (2002)

AARP guidelines (2004)

SilverWeb guidelines (2005)

Webcredible guidelines (2006)

The following sub-sections will provide an overview of each of these sets of guidelines.

2.6.1   SPRY Foundation guidelines (1999)

Holt and Komlos-Weimer (1999) published guidelines through the SPRY foundation entitled "Older adults and the web: a guide for web site creators". The guidelines were developed as an outcome of a conference on "Older Adults, Health Care Information, and the World Wide Web" held in March 25-26, 1999. The authors stated that the guidelines were developed by combining research with practical experience from experts. However, no research evidence was provided in the publication. There was also no evidence to show which recommendations were supported by research or which recommendations were suggested by experts. Many recommendations are very similar to the WCAG 1.0 checkpoints. Later web design guidelines for older people seem to be based on the SPRY Foundation guidelines, especially Holt (2000).

2.6.2   Holt (2000) guidelines

Holt (2000) presented the guidelines "Create Senior-Friendly Web sites" in the journal of the Centre for Medicare Education. No information was provided about how these guidelines were derived. However, when considered in detail, it was appears that these guidelines are very similar to the SPRY Foundation guidelines (1999) and there are many places in these guidelines with the same wording. However, in these guidelines there are some different recommendations. For example 14 point size text is recommended instead of 12 to 14 point. It seems that the SPRY guidelines were updated by Holt (2000), but no specific evidence for such updates were provided.

2.6.3   Zhao (2001) guidelines

The guidelines proposed by Zhao (2001) were named "Universal Usability Web Design Guidelines for the Elderly (Age 65 and Older)". No information was provided about

how these guidelines were derived, but it seems that the guidelines were developed from reviewing the guidelines from many sources and merging them.

Zhao (2001) stated that the recommendations in the guidelines relate to hardware, software and input devices can be provided to enhance accessibility. However, the guidelines suggested by Zhao (2001) were quite similar to previous guidelines, such as WCAG 1.0, and the AgeLight guidelines (see below) which were released at nearly same time.

2.6.4   AgeLight (2001) guidelines

AgeLight (2001) published the "Interface design guidelines for users of all ages" in 2001. It is claimed that this set of guidelines was created from "dozens of focus groups, feedback from users, cooperation with people in the fields of usability, human factors, and aging as individuals and organizations". However, no specific evidence from these different sources is provided, nor information about how they contributed to the development of the guidelines.

Although AgeLight's guidelines (2001) stated to be concerned with accessibility specifically for older people, some of the guidelines are general accessibility guidelines, such as providing a text version of a web site which would apparently be suitable for blind readers, but would certainly not be used by older people. In addition, the AgeLight guidelines are concerned with vision loss more than other impairments, on this issue they provide a lot of information, examples, and guidelines.

2.6.5   National Institute of Aging / National Library of Medicine (2002) guidelines

Hudes and Linberg (2002) developed guidelines under the auspices of the National Institute of Aging and National Library of Medicine (NIA/NLM), integrating other guidelines with many research results. The aim of the guidelines was to make "senior friendly" web sites. Hudes and Linberg (2002) did not provide information that how the guidelines were derived. However, Morrell (2005) explained that these guidelines were developed from reviewing the research in the fields of cognition and ageing, perception and ageing, human factors and ageing. Book chapters, book, journals and presentations in conferences also included. However, I have found that some of the guidelines are taken from research about text presentation on print media.

The NIA/NLM guidelines are divided into three groups: recommendations relating to aged-related declines in vision, recommendations relating to aged-related declines in cognition, and other issues that are important to take into account in the design of web sites for older people.

2.6.6   American Association of Retired Persons (AARP) (2004) guidelines

Redish and Chisnell (2004) reviewed documents about web site design for older people which were published between January 2000 and September 2004 and used this information to create guidelines for web site design for older people under the name of American Association of Retired Persons (AARP).

Most of the guidelines in this set of guidelines repeat the recommendations from previous sets. However, the guidelines were not sufficiently clear and conflict within their guidelines. For example, the guideline about links in website has two different recommendations. One recommends having multiple links which lead to same content as this will increase opportunities for older people in reaching target information. However another recommends that older people read slower on pages which have a high number of distractors. A large number of links lead to a high number of distractors, so these recommendations are in conflict with the another.

In addition, AARP does not suggest the appropriate number of the links, so it is not clear and difficult to apply.

2.6.7   SilverWeb guidelines (2005, 2007)

Kurniawan and Zaphiris (2005) and Zaphiris, Kurniawan, and Ghiawadwala (2007) reviewed more than 100 research papers about HCI and ageing to derive guidelines about web design for older people. From this review, they created a set of 52 guidelines. They used card sorting by 40 postgraduate students and a focus group of five HCI experts to classify these guidelines into ones related to: vision (changes in static acuity, dynamic acuity, contrast sensitivity, colour sensitivity, sensitivity to glare, decrease in visual field, and decrease in processing visual information), psychomotor abilities, attention (changes in selective and divided attention), memory and learning, intelligence and expertise. Each guideline was backed up with at least one published piece of literature or study. They named their guidelines the "SilverWeb Guidelines".

These guidelines have a focus on older people. However, rather than trying to establish whether the guidelines have a good empirical evidence basis in terms of their effects

on the performance and preferences of older people when using the web, the researchers used younger people to categorise the guidelines and had them checked by HCI experts. They also asked 16 older people to rate the usefulness of each guideline in relation to an evaluation of two websites. These older people were not asked to do tasks on the websites, so it is not clear on what basis they were rating usefulness or how these ratings provide any appropriate validation of the guidelines.

2.6.8   Webcredible guidelines (2006)

The web usability consultancy company Webcredible created a set of guidelines for older adults (Fidgeon, 2006) by conducting 40 minute talk-aloud usability test sessions, 8 with older participants (over the age of 65) and 8 with young participants (under the age of 40). Participants were asked to find information from a range of government web sites. This is a very small sample on which to base guidelines, being from only 8 older people and based on one particular kind of website.

2.6.9 Comparison of the eight sets of web design guidelines for older people

The eight sets of web design guidelines for older people as mentioned in Section 2.6.8 provide recommendations by using the difference names or categories, then it is hard to understand and compare each recommendation in details. This section tries to organise the recommendations from different web design guidelines to be a set of recommendation.

Morrell (2005) in reviewing research relevant to web design guidelines for older people divided the problem areas into three groups: vision, cognitive abilities and other problem areas for older people, as shown in Table 2.7. I have used this classification to group all the recommendations in the eight sets of web design guidelines reviewed above. Entries in Table 2.7 with an asterisk are those from Morrell (2005) and those without are additional entries that I have added to cover recommendations not reviewed by Morrell. The figures in brackets are the number of sets of guidelines which provided the recommendation.

From table 2.7, the first group of the guidelines called 'guidelines related to abilities in vision'. There are 11 guidelines in this group. Most of the guidelines are related to text presentation on the web. Another group called 'guidelines related with abilities in cognition' has 11 guidelines. Most of the guidelines in this group related to recommendations about how to make information on the web easy to understand. The last group called 'other issues in web design guidelines for older people'. There are 16

guidelines in this group. Most of them are related to structure, navigation, and other dimensions for creating web pages.

For my research programme I decided to concentrate on those guidelines related to vision loss. This is because for older people, vision loss has the greatest impact on their use of the web (AgeLight, 2001) and is the reason for many difficulties in their use of technology (Dickinson et al, 2005). Moreover, as can be seen from Table 2.8, all the guidelines related to vision loss provide recommendations on the presentation of text. Although the web has now become much more multimedia, and visual presentation needs to cover other aspects such as images, video and animation, text is the most basic visual presentation on the web, and we need to make sure older users can access it easily.

Table 2.8: Web design guidelines related to changes in vision, cognition, and other issues in older people (adapted from Morrell, 2005) with number of mentions across the 8 sets of guidelines discussed

Guidelines related to abilities in vision	Guidelines related with abilities in cognition	Other issues in web design guidelines for older people
- Line spacing (5)	- style of writing* (4)	- Navigation and page location* (7)
- Text justification* (6)	- Phrasing* (2)	- use of mouse* (4)
- Font type* (7)	- Simplicity* (5)	- Forward and backward navigation* (2)
- Font size* (7)	- Illustrations and photographs* (5)	- Consistent layouts* (6)
- Type weight* (4)	- Animation, audio and video* (7)	- Style and size of icons and buttons* (8)
- Capital and lowercase letters*(4)	- Text alternatives* (6)	- Pull-down menus* (5)
- Kerning (3)	- other issues to consider organization and repetition* (3)	- Site maps* (4)
- Backgrounds* (8)	- avoid technical term (4)	- Scrolling* (5)
- Colour* (7)	- break content to short section (2)	- Length of page (3)
- White space (3)	- provide fewer choices	- Opening new browser (3)
- Length of line (1)		- Hypertext links (7)
		- Search engine and search capability (3)
		- Online help tutorial, instruction,

	to user (1) - Provide ample time to read content (1)	About us and feedback (8) - Support user control and freedom (2) - Table and frame (2) - Usability testing (6) - others (Design for Internet appliances (1), Provide content in HTML as much possible (1), provide date stamping (1), archive old article (1), design for slow modem (1), design for visual impairment software (1))

N.B. Entries with asterisks indicate those taken from Morrell (2005).

In the next section, I will present the specific details about the various recommendations of web design guidelines for older people in relation to vision loss and text presentation.

2.7 Specific recommendations in web design guidelines for older people related to presentation of text on web pages

Table 2.9: Recommendations on line spacing from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on line spacing
SPRY Foundation (1999)	Even average or default leading may not be sufficient for those with this problem [blocks of text appear crowed], which occurs frequently in older adults. This is easily remedied, however, by increasing the leading by even a small amount (1 or 2 points).
Holt (2000)	Older adults may have more trouble reading pages that are single-spaced rather than double-spaced. In particular, if bold type is used, it is better to add the extra space to improve both the ease and the speed of reading. An alternative is to format your paragraphs at 1½ spaces, or add a few extra points of space between lines; this "airs

	out" your text without using as much screen space.
Zhao (2001)	Not mentioned
AgeLight (2001)	The space between each line of text is required. Typically, the leading specified is 2 points larger than the typeface. Tight leading may cause legibility problems. Loose leading may bring benefits by creating breathing room and improve readability.
NIA/NLM (2002)	Double space all body text.
AARP (2004)	Not mentioned
SilverWeb (2005, 2007)	There should be spacing between the lines.
Webcredible (2006)	Not mentioned

Table 2.9 shows that five of the sets of guidelines reviewed make recommendations about line spacing: SPRY Foundation (Holt and Komlos-Weimer, 1999), Holt (2000), Agelight (2001) and NIA/NLM (2002) provide quite specific recommendations, whereas SilverWeb (2005, 2007) provide very general recommendations. The other guidelines reviewed, Zhao (2001), AARP (2004) and WebCredible (2006), do not mention line spacing at all. Even the specific recommendations are diverse, with one recommending 1.5 spacing and another double spacing. Thus, there is no consensus amongst the recommendations about line spacing across the various web design guidelines for older adults.

Table 2.10 shows that six of the sets of web design guidelines provide recommendations on text justification. In this case, six of the eight sets of guidelines make the same recommendation, for left-justified text. The final two sets of guidelines, AARP (2004) and WebCredible (2006), do not mention text justification.

Table 2.10: Recommendations on text justification from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on text justification
SPRY Foundation (1999)	Most older adults prefer left justified text, where the text lines up along the left margin, and find it easiest to read.
Holt (2000)	Justification refers to how words are placed within the page margins. Type that is centered is fine for headings

	but can be difficult to read in paragraph form. Full justification (spread evenly between the margins) adds extra spaces or reduces spaces between letters and words. On shorter width text where there is less to change, full justification can make reading uncomfortable.
Zhao (2001)	Left-hand justification offers the highest level of readability. Center justification other than for a title, should be avoided.
AgeLight (2001)	Left-hand alignment offers a high level of readability as compared to justification.
NIA/NLM (2002)	Left justified text is optimal for older adults.
AARP (2004)	Not mentioned
SilverWeb (2005, 2007)	Text should be left justified.
Webcredible (2006)	Not mentioned

Table 2.11 shows that there are seven sets of web design guidelines which provide recommendations about font type. Four sets recommend using a Sans Serif font (Holt, 2000; NIA/NLM, 2002; AARP, 2004; SilverWeb, 2005, 2007) and another two make a more general recommendation about using a font for familiarity and legibility (Agelight, 2001; Zhao, 2001), a somewhat circular recommendation. Finally, two sets of guidelines either make no recommendation about font (WebCredible, 2006) or sit on the fence between Serif and Sans Serif fonts (SPRY Foundation, 1999). Thus, the recommendation for using a sans serif font is the most frequently made, but is only recommended by half the sets of guidelines.

Table 2.11: Recommendations on font type from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on font type
SPRY Foundation (1999)	Traditional design wisdom holds that Serif fonts in a mix of upper and lowercase letters (known as sentence case) are the most readable for a block of text. However, there is some evidence that Sans Serif fonts such as Helvetica are the most readable for older adults.
Holt (2000)	For publication on the Web, Sans Serif faces (those

	without extra strokes in the letters) are generally considered easier to read. (Note: this differs from printed materials where Serif fonts are considered to be easier to read.)
Zhao (2001)	For print applications, Serif typefaces are more legible because the Serif adds differentiation between letter forms, yet on lower resolution and small monitors, this may not always be true. Choose fonts based on their legibility, and avoid using several types of fonts mixed together or very narrow or decorative fonts.
AgeLight (2001)	Choose typefaces based on their familiarity and legibility.
NIA/NLM (2002)	Use a Sans Serif typeface, such as Helvetica, that is not condensed. Avoid the use of serif, novelty, and display typefaces.
AARP (2004)	San Serif is recommended. Bernard et al. (2001) recommended Serif for speed and San Serif for preference.
SilverWeb (2005, 2007)	Use San Serif type font i.e., Helvetica, Arial. Avoid other fancy font types.
Webcredible (2006)	Not mentioned

Table 2.12 shows the recommendations on font size for seven sets of design guidelines; six sets of guidelines make recommendations for between 12 and 14 point size, with somewhat different wordings. One set of guidelines merely warns that the smaller the font, the more important it is to make each letter or line distinguishable. Thus, the recommendations on font size is generally between 12 and 14 point.

Table 2.12: Recommendation on font size from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendations on font size
SPRY Foundation (1999)	Most older adults prefer a font size somewhere between 12 point and 14 point for blocks of text, depending on the typeface involved. Headings should be enough larger than the body text to be distinguishable as well as readable. An

	18pt heading would be a good choice to accompany a 14pt block of text.
Holt (2000)	The smaller the type size, the more critical it becomes to have each letter or line distinguishable from the ones around it.
Zhao (2001)	For most seniors, 12 to 14 point fonts are recommended for body [text] while headlines and titles are typically two points larger. Those with partial sight may require a 16 point font or above
AgeLight (2001)	12 - 14 point are recommended font sizes for copy while headlines and titles are typically two points larger.
NIA/NLM (2002)	Use 12 point or 14 point type size for body text.
AARP (2004)	Use at least 12 point, Some suggest 14 point for body text and heading should be 18 and 24 point. Bernard et al. (2001) said older people read 14 point faster than 12 point.
SilverWeb (2005) (2007)	Use 12 -14 point size.
Webcredible (2006)	Not mentioned

Table 2.13 shows recommendations on type weight from four sets of web design guidelines. These recommendations cover two aspects of type weight – that for the main text and the use of bold text. Two sets of guidelines cover the main text aspect and both recommend medium weight (Agelight, 2001; Zhao, 2001). The other two sets of guidelines cover the use of bold and both recommend using bold only for emphasis (NIA/NLM, 2002; SPRY, 1999). .

Table 2.13: Recommendations on type weight from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on type weight
SPRY Foundation (1999)	Type needs to be intense enough to be clearly read, but not so bold as to be hard to distinguish. Medium weight types frequently provide a good contrast with the background without becoming too intense.

Holt (2000)	Not mentioned
Zhao (2001)	Many typefaces are available in light, narrow, bold, or extra bold. While boldfaced text may appear larger, its readability is decreased. Use bold only to emphasize a title or a key word.
AgeLight (2001)	Boldfaced text appears larger, readability may decrease. Limiting the use of bold to emphasize a title or a key word is recommended.
NIA/NLM (2002)	Use medium or bold face type.
AARP (2004)	Not mentioned
SilverWeb (2005) (2007)	Not mentioned
Webcredible (2006)	Not mentioned

Table 2.14 shows recommendations on the use of capital and lowercase letters from four sets of web design guidelines. All four recommend avoiding text in all capitals, or restricting it to keywords and titles (Agelight, 2001; Zhao, 2001; NIA/NLM, 2002; SilverWeb, 2005, 2007). In addition, two sets of guidelines recommend using a capital letter for the first letter of a heading or title (Zhao, 2001; NIA/NLM, 2002).

Table 2.14: Recommendations on capital and lowercase letters from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on use of capital and lowercase letters
SPRY Foundation (1999)	Not mentioned
Holt (2000)	Not mentioned
Zhao (2001)	Using all capital letters decreases readability. While sometimes used for design purposes, it tends to lead to higher levels of eyestrain and eye fatigue because there is too little differentiation between the letters, and the eye does not get a visual breather. At best, only use capital letters for key words or titles. Capitalize the first letter of each word in a heading instead of all of it, although bold type is recommended as a more effective alternative.

AgeLight (2001)	The use of all caps tends to lead to higher levels of eye fatigue because of little differentiation between the letters. As an alternative, consider using bold or capitalize the first letter of each word in a heading. This provides contrast from the body copy, will increased readability.
NIA/NLM (2002)	Present body text in upper and lowercase letters. Use all capital letters and italics in headlines only.
AARP (2004)	Not mentioned
SilverWeb (2005) (2007)	Main body of the text should be in sentence case and not all capital letters.
Webcredible (2006)	Not mentioned

Table 2.15 shows recommendations on kerning[1] from three sets of web design guidelines. Kerning is the process of adjusting the spacing between characters in a proportional font, usually to achieve a visually pleasing result

The recommendations on kerning are not specific; two sets of guidelines recommend avoiding kerning or too much kerning. without specifying what "too much" is (SPRY, 1999; Holt, 2000). One set of guidelines (Agelight, 2001) mentions kerning without making a recommendation about it.

Table 2.15: Recommendations on kerning from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on kerning
SPRY Foundation (1999)	Designers should avoid condensed typefaces and reducing the kerning (space between letters), as this can also make text harder to read.
Holt (2000)	Using condensed type squeezes the letters together and makes them harder to read. While effective for squeezing more copy onto a page, too much of it together can appear blurry and cluttered.
Zhao (2001)	Not mentioned

---

[1] http://en.wikipedia.org/wiki/Kerning.

AgeLight (2001)	A designer can specify tight, regular or loose letter spacing to be applied throughout a design or style sheet. Specific adjustments can be made between letters to enhance legibility. For example, the space between a capital A and lower case letters often needs to be kerned to make the space smaller.
NIA/NLM (2002)	Not mentioned
AARP (2004)	Not mentioned
SilverWeb (2005, 2007)	Not mentioned
Webcredible (2006)	Not mentioned

Table 2.16 shows recommendations on text and background colour contrast from the eight sets of web design guidelines. Six sets of guidelines recommend using dark text on a light background (SPRY, 1999; Holt, 2000; Zhao, 2001; Agelight, 2001; NIA/NLM, 2002; Webcredible, 2006). Two sets of guidelines also mention that "reverse contrast" can be used, a light text on a dark background (Holt, 2000; NIA/NLM, 2002). Two sets of guidelines mentions only that strong contrast between text and background should be used (AARP, 2004; SilverWeb, 2005, 2007). Two sets of guidelines also mention that off-white rather than pure white is a better background (SilverWeb, 2005, 2007; Webcredible, 2006). So the consensus across the guidelines is for dark text on a light background, but possibly off-white rather than pure white.

Table 2.16: Recommendations on contrast between text and background from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on text and background colour
SPRY Foundation (1999)	A good background … should contrast with the content of a web site in all three areas. Traditional designs typically feature dark text on a light background, but there is some evidence that older adults find a light text on a dark background to be very readable. As long as the contrast is strong, either is acceptable.
Holt (2000)	Contrast between background and text is important. Usually this means having dark type or graphics against a light background (though sometimes the opposite can

	work fine with sufficient contrast).
Zhao (2001)	To maximize contrast, always use dark types on light or white backgrounds, exaggerate lightness differences between foreground and background colors, and avoid using colors of similar lightness adjacent to one another, even if they differ in saturation or hue. It is a good practice to choose dark colors with hues from the bottom half of the color wheel against light colors from the top half of the circle. Avoid contrasting light colors from the bottom half against dark colors from the top half. Also, be aware that people with color deficits will see less contrast between colors. So it helps to even lighten light colors and darken dark colors.
AgeLight (2001)	Contrast between foreground and background colours. As a rule use dark type on light or white backgrounds.
NIA/NLM (2002)	Use dark type or graphics against a light background, or white lettering on a black or dark-coloured background.
AARP (2004)	When contrast between type and background is low, users can also suffer "disability glare," causing them to "lose" letters in text passages. High contrast also makes it easier for older adults to remember what they've seen and read and to make inferences from text.
SilverWeb (2005, 2007)	Background screens should not be pure white or change rapidly in brightness between screens. Also, a high contrast between the foreground and background should exist, for example, coloured text on coloured backgrounds should be avoided.
Webcredible (2006)	Always use high contrast to display text e.g. black text on an off-white background (N.B. using an off-white background is preferable to white because it reduces the chances of eyestrain for people who are slow readers).

Table 2.17 shows recommendations on use of colour in websites in general and website "wallpaper" or patterned backgrounds in particular from six sets of web design guidelines. Five sets of guidelines recommend avoiding colour combinations in the blue/yellow, blue/green or blue/yellow/green colour space (SPRY, 1999; Holt, 2000;

Agelight, 2001; NIA/NLM, 2002; SilverWeb, 2005, 2007).   Three sets of guidelines recommend avoiding colour combinations in the red/green colour space (SPRY, 1999; Holt, 2000; Agelight, 2001).   Two sets of guidelines recommend avoiding bight, neon colours (Holt, 2000; Zhao, 2001). Two sets of guidelines make a positive recommendation, to use complementary colours from opposite sides of the colour wheel (Agelight, 2001; Zhao, 2001).   And finally three sets of guidelines recommend avoiding patterned backgrounds that compete with the text (SPRY, 1999; Agelight, 2001; NIA/NLM, 2002), although light backgrounds may be useful.

Thus the range of recommendations given by the guidelines on colour combinations to avoid is rather extensive, and web developers may feel quite limited by them. However, at least some of the sets of guidelines offer the positive recommendation of using complementary colours.

Table 2.17: Recommendations on general use of colour from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendations on colour
SPRY Foundation (1999)	A good background or wallpaper can enhance the content of a page and draw the user's attention to a particular segment or graphic; a bad background can upstage the content and make it impossible to understand. Contrast, to be most effective, needs to occur in three different areas: hue, saturation and lightness. Red is demised by its complimentary color being mixed in (green).
Holt (2000)	Your choice of color matters a great deal. Decreased sensitivity to color can make distinguishing between certain colors difficult for seniors, particularly red/green and blue/yellow combinations. Bright neon colors may also become annoying.
Zhao (2001)	Choose complementary colors: The color wheel is a tool that arranges the colors of the spectrum by hue. It is recommended to choose colors from opposite sides of the color wheel, e.g. when choosing a primary color such as blue, its complementary color would be orange. Avoid some colors: Colors that are exceptionally bright,

	fluorescent, or vibrant can have edges that appear to blur and create after-images, which tire the eyes. For example, yellow text is very difficult to read. Alight type color on a dark background can cause the type to appear to "close in itself". Avoid short wavelength and blue-green regions.
AgeLight (2001)	Use colours which opposite side of colour wheel. Avoid combinations of blue and yellow or red and green as many users have some degree of colour deficiency or colour blindness in these areas. Using any background patterns including watermarks or embossed logos generally are distracting and interfere with readability. As an alternative, a light complementary background colour can be applied.
NIA/NLM (2002)	Avoid yellow and blue and green in close proximity. These colours and juxtapositions are difficult for some older adults to discriminate. Ensure that text and graphics are understandable when viewed on a black and white monitor. Avoid patterned backgrounds.
AARP (2004)	Not mentioned
SilverWeb (2005, 2007)	Colours should be used conservatively. Blue and green tones should be avoided. Content should not all be in colour alone (colour here is denoted by all colours other than black and white).
Webcredible (2006)	Not mentioned

Table 2.18 shows that there is only one set of guidelines reviewed above which make recommendations about line length: SilverWeb (2005, 2007) state that line length should be should be short, but there is no information about what short actually means.

Table 2.18: Recommendations on line length from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on length of line
SPRY Foundation (1999)	Not mentioned
Holt (2000)	Not mentioned
Zhao (2001)	Not mentioned
AgeLight (2001)	Not mentioned
NIA/NLM (2002)	Not mentioned
AARP (2004)	Not mentioned
SilverWeb (2005, 2007)	Text line should be short in length
Webcredible (2006)	Not mentioned

Table 2.19 shows recommendations on white space from the eight sets of web design guidelines. Two sets of guidelines recommend that small blocks of text on large areas of white space increase readability (Agelight, 2001; Zhao, 2001). Another set of guidelines points out the important of white space (AARP, 2004). However, no set of web design guidelines provide information about the appropriate proportion of white space to text.

Table 2.19: Recommendations on white space from the eight sets of guidelines on web design for older people

Web design guidelines	Recommendation on white space
SPRY Foundation (1999)	Not mentioned
Holt (2000)	Not mentioned
Zhao (2001)	Large areas of white space and small blocks of text increase readability, making pages cleaner looking and easier to navigate. If possible, use short text or lists to paragraphs of text. However, larger blank space causes larger pages that mean more scrolling.
AgeLight (2001)	Large areas of white space and small blocks of text increase readability.

NIA/NLM (2002)	Not mentioned
AARP (2004)	White space is also important within the main sections of the page. Incorporating appropriate white space and line spacing facilitates clicking links, icons, and other small targets.
SilverWeb (2005, 2007)	Not mentioned
Webcredible (2006)	Not mentioned

As a result of this analysis, some of the recommendations from the sets of web design guidelines related to vision loss were selected to empirically investigate their effects with older people. Firstly, the recommendations related to text presentation are the main interest of the current programme of research. A further reason for making this choice is most of the content on the web presenting as text, then the large number of web design guidelines make recommendations in this area. In addition, the areas in which different sets of guidelines make different recommendations have been selected, as it is particularly important to establish which is the most appropriate recommendation with empirical research with older people. Finally, recommendations which cannot apply to both the Thai and Latin alphabets were not selected, as in my programme of research I wished to compare older people's use in both the Thai and Latin alphabets. Thus, line spacing, text justification, font type, font size, text and background colour were selected to investigate in the experiments which are presented in Chapter 3 to 6.

The next section presents previous research on the recommendations that I selected for investigation.

2.8 Previous research relevant to specific recommendations made by web design guidelines for older adults

The following sections will review previous research relevant to the five areas of web design for older adults chosen for investigation in this programme of research: line spacing, text justification, font type, font size and colour of text and background. The review will be restricted to research which has investigated reading from screens, and will not cover earlier research that investigated reading from paper or comparing paper to screens. However, it will cover research about younger readers as well as older readers, as there is so little research about older readers and research on younger

readers provides a useful baseline from which to work. This research helped formulate not only my research questions, but also the methodology to be used.

2.8.1   Research on the effects of line spacing in reading from screens

The earliest study found on the effects of line spacing on reading from a computer screen was by Kolers, Duchnicky and Ferguson (1981). They studied readability from Cathode Ray Tube (CRT) displays using an eye tracking methodology. They investigated single and double line spacing. There were 20 participants, most of them female university students (no age information was given). Each participant was asked to read 20 texts with 20 conditions in random order. Each text was 300 words long. The dependent variables were total number of fixations, number of fixations per line, number of words per fixations, fixation rate, fixation duration, and total reading time. For line spacing there were significant effects for total number of fixations, number of fixations per line, number of words per fixations, and total time. Double line spacing reduced the total number of fixations by about 3%, the number of fixations per line also by 3%, but it increased the number of words per fixation by about 4% and most importantly the total reading time by 2%. The authors concluded that such small differences, while statistically significant, were of little practical importance.

This study only studied two levels of line spacing, single and double spacing. It also used CRT displays which are now a completely outdated technology. In addition, the participants were only younger female adults who were university students. Thus the results may not be relevant to the performance of older adults with modern screen technologies. However, this study is the earliest study found about reading text from a screen.

Kruk and Muter (1984) conducted a series of three experiments about reading continuous text from a a 30.5 cm (diagonal) green monochrome video monitor (Amdek 100G). Their third experiment is relevant here, as the independent variables included two levels of line spacing: single and double spacing. The dependent variables were reading speed and comprehension scores. The participants were 12 university students (no ages were given). Participants were asked to read four sets of materials for 5 minutes each. There was a significant main effect for line spacing: reading text with single line spacing was 10.9% slower than with double line spacing. There were no significant interactions with other variables in the study. Again, this study was with a

now outdated display technology and with young readers, so may not be relevant to older people with current screen technologies.

Chan and Lee (2005) investigated the effect of text presentation factors on reading from a 15 inches colour cathode ray tube (CRT) screen in Chinese. The independent variables included two levels of line spacing (single and double spacing). The dependent variables were reading time, comprehension scores, and preference ratings. 72 university students participated in the study, aged from 19 to 24 years. Line spacing had a significant effect on reading speed: participants read faster with double line spacing over single line spacing. However, line spacing had no effect on comprehension scores. To elicit preference measures, participants were asked to rate four attributes: reading comfort, reading ease, reading fatigue, and overall preference on 9 point Likert items. Double line spacing was rated significantly more comfortable, more easy, less fatiguing and overall more preferred than single line spacing.

This study showed a significant effect of line spacing both performance and preference for Chinese text. However, the study investigated only two levels of line spacing: single and double line spacing. In addition, the participants were university students so only represent very young adults.

Ling and Schaik (2007) investigated the influence of line spacing, amongst other variables, using a visual search task on web pages running on personal computers (Intel Pentium, 333 MHz, 64Mb RAM, Microsoft NT4 operating system, 14 inch monitors). They investigated spacing single, 1.5 and double line spacing. The dependent variables were accuracy, speed, and aesthetic appeal, as measured by four items from the scale developed by Tractinsky, Katz and Ikar (2000). Participants were undergraduate students, a majority (77%) were 25 or under; the rest were aged between 26 and 50. The participants searched for a target hypertext link amongst five hypertext links in a screen of text. The participants completed the tasks as quickly but as accurately as possible. After completing the search tasks, they were presented with all three possible pairs of line spacing. Participants then were asked to choose which of the pair they preferred. The same procedure was used with text justification pairs. Participants performed significantly more accurately and with faster reaction times o with increasingly wider line spacing. In addition, measures of aesthetics and preference also increased significantly with wider line spacings.

This study used a modern display to present the text materials, three levels of line spacing and the participants were not all very young. However, they were still all under

50 years of age, and 75% were under 26 years of age, so it is not clear that these interesting results would generalize to older adults.

Chan, Tsang, and Ng (2014) investigated the effects of line spacing and other text presentation variables on reading in Chinese on 17" liquid crystal display monitor, . Three levels of line spacing were used: single, 1.5, and double line spacing. 39 undergraduate students participated in the study, aged from 21 to 26 years. The dependent variables were proofreading time, typographic error detection rate, and amount of scrolling. There was a significant effect of line spacing on proofreading time: proofreading time increased significantly with wider line spacing. Line spacing also had significant effect on typographic error detection rate: 1.5 and double line spacing produced significantly higher detection rate than single line spacing but there was no difference between 1.5 and double line spacing. However, line spacing did not affect the amount of scrolling.

This study investigated effect of line spacing with three levels of line spacing (single, 1.5, double line spacing) and produced interesting results for Chinese. However, this study used only younger adults.

Only one paper could be found which studied the effects of line spacing on reading from screens with older people with normally aging sight and this was on mobile phones rather than computer screens. Wang, Sato, Rau, Fujimura, Gao and Asano (2009) investigated the effects of line spacing, amongst other variables, on reading in Chinese from mobile phone screens. They used NEC N6305 mobile phones with a screen size of 30 x 38 mm. The resolution was 176 x 220 pixels. One independent variable was four levels of line spacing: 2, 4, 6 and 8 pixels. Tasks were reading and visual search. Dependent variables were reading performance (time and errors), text readability, visual fatigue, and preferences. There were no significant differences in the performance measures, however there were significant differences in the preference measures. Perception of readability increased with increasing line spacing.

In summary, although there is a small body of research on the effects of line spacing on both participants' reading performance and preferences when reading from screens, this is hardly comprehensive, and there is very little research that investigates the these effects for older readers.

2.8.2 Previous research on the effects of text justification in reading from screens

Only one paper could be found which investigated the effects on text justification in reading from screens, Ling and van Schaik (2007), already discussed in relation to line spacing (see section 2.8.1). They found an interesting effect in that participants performed better with Left-aligned text, but preferred justified text. No studies could be found which investigated the effects of text justification in reading from screens with older adults.

2.8.3 Previous research on the effects of font type in reading from screens

Tullis, Boynton and Hersh (1995) investigated the readability of font types for Microsoft Windows applications on a NEC 5FG 15" monitor running in 1024 x 768 (Small Fonts) resolution. The independent variables included four levels of Font Type: Arial, MS Sans Serif, MS Serif and Small Font. The dependent variables were reading time, accuracy, and preferences. There were 15 volunteer participants aged between 27 and 45 years. The authors did not provide more details about participants. The participants were asked to read 48 combinations of text and find typographical errors within each text. Unfortunately the researchers did not analyse their data with Font Type as a separate variable in the analysis of variance, which they could have, but used an independent variable of Font Type/Size. Thus one cannot extract the effects of Font Type separately. There were significant differences between font/size combination in reading time, accuracy, and preference. The researchers concluded that most of the fonts from size 8.25 point to 9.75 point produced acceptable reading time and accuracy, except MS Serif 8.25 point. The most preferred font/size combinations were Arial 9.75 point and MS Sans Serif 9.75 point. Another suggestion was avoid to use Arial 7.5 point, Small Font 6.0 point, and Small Font 6.75 point.

Boyarski, Neuwirth, Forlizzi and Regli (1998) conducted three studies to investigate the effects of font type on reading. The second study is relevant here as it compared reading from a 17" Sony Trinitron multiscan 17seII monitor with Georgia (a serif font) and Verdana (a san serif font). The dependent variables were comprehension scores, reading time, effective reading speed (score/time), and reader preferences. There were 16 participants, students and university staff, aged 20 – 53 years. There was no significant difference in reading time and effective reading speed. In terms of preference, participants significantly preferred Verdana over Georgia in relation to ease

of reading but there was no significant difference in relation to most pleasing to read and most sharp.

Bernard, Mills, Peterson and Storrer (2001) investigated the effect of 12 popular online fonts presented on a Pentium II based PC computer with a 60 Hz, 96dpi 17 inch monitor. The fonts comprised five Sans Serif fonts: Agency, Arial, Comic, Tahoma, and Verdana; five Serif fonts: Courier, Georgia, Goudy, Century Schoolbook, and Times New Roman; and two ornate fonts: Bradley, and Monotype Corsiva. All the fonts were used in 12 point except for Agency which was 14 point in order to have the same physical size as the other fonts. The dependent variables were reading efficiency (the percentage of accurately detected substituted words in the passages, divided by the time taken to read the passages), reading time, and participants' perceptions of the fonts on six dimensions (legibility, personality, elegance, youth and fun, business-like, and general preference). There were 22 participants aged 20 to 44 years, mean age 25 years. However, the researchers did not provide other information about participants. Each participant was asked to read 12 passages and find substituted words (words which had been altered in the text, substituted with a grammatically incorrect but similar looking word, for example "fake" for "cake"). There was no significant effect of Font Type on reading efficiency. However, there was significant on reading time, Tahoma font was read significantly faster than only the Corsiva font. There was no other significant different on reading time. On perception of font legibility, Courier, Comic, Georgia, Verdana, and Times New Roman were perceived as being significantly more legible than Agency, Bradley, Goudy, and Corsiva font types. On the perception of personality of the font, Bradley was significantly higher in rating of personality than Courier, Tahoma, Goudy, Schoolbook, or Times New Roman while Times New Roman font received significantly lower perceptions of personality than Bradley, Comic, and Corsiva. On perception of elegance, Bradley was pereceived as significantly more elegant than Agency and Courier while Corsiva was significantly more elegant than all other font types except Bradley. On the perception of youthful and fun, Comic was perceived as significantly more Youthful and Fun than Arial, Agency, Courier, Schoolbook, Goudy, and Times New Roman, while Times was perceived as being significantly less Youthful & Fun than Georgia, Verdana, and Comic. On perception of business-like, Times New Roman and Courier were significantly perceived as being more business-like than all font types except Tahoma, Verdana, Georgia, and Schoolbook. On general preference, Arial, Comic, Tahoma, Verdana, Courier, Georgia, and Schoolbook were significantly preferred over the other font types, with Verdana the first, and Arial and Comic the second preference choices.

Bernard, Liao, and Mills (2001a, 2001b) conducted a study which investigated the effect of four popular font types: Time New Roman, Georgia, Arial and Verdana, presented on a Pentium II based PC computer with a 60 Hz, 96dpi 15 inch monitor. The independent variables were organized into two levels of font type (serif fonts and san serif fonts). Two levels of Font size were included as another independent variable, but the results related to font size are presented in Section 2.8.4, below. There were 27 older participants, age range 62 to 83 years, who had experienced with reading on screen. The participants were asked to find substituted words (see discussion of previous paper by Bernard et al 2001) in the texts. The dependent variables were reading efficiency (the percentage of accurately detected substituted words in the passages, divided by the time taken to read the passages), reading time, and rating of preference. Font type was no effect on both reading efficiency and reading time. However, the two most preferred combinations were all Sans Serif fonts. The authors concluded that the selection of font type for older computer users based on user preference was san serif font types.

Bernard, Lida, Roley, Hackler and Janzen (2002) conducted a study which included eight popular online fonts (Century Schoolbook, Courier New, Georgia, Times New Roman, Arial, Comic Sans MS, Tahoma, and Verdana) presented on a Pentium II based PC computer with a 60 Hz, 96dpi 17 inch monitor. The dependent variables were reading efficiency (as described in summary of research from Bernard, Liao, and Mills (2001a, 2001b), above), reading time, perception of font legibility, font attractiveness, and general preference. There were 60 participants aged 18 to 55 years, with a mean of age of 24 years. Participants were again asked to find substituted words. There was no significant of font type on reading efficiency. However, there were significant effects on reading time. Courier New, Century Schoolbook, and Georgia were read significantly more slowly than Times New Roman and Arial. On perceived legibility, Arial and Courier were considered the most legible fonts, whereas Comic was perceived as the least legible font. On perceived attractiveness, Georgia was perceived as being significantly more attractive than Arial, Courier, and Comic, while Times New Roman was perceived as significantly more attractive than Courier. On font preference, Times New Roman was significantly less preferred to all fonts except Schoolbook. Schoolbook was significantly less preferred in comparison to Verdana. Overall, Verdana was the most preferred font, while Times was the least preferred font.

Bernard, Chaparro, Mills and Halcomb (2003, a preliminary version of this study was presented in Bernard and Mills, 2000) investigated the effect of two popular font types

Times New Roman and Arial, amongst other variables, presented on a Pentium II based PC computer with a 60 Hz, 96dpi 17 inch monitor. 35 participants, aged 17 to 47 with a mean age of 25 years took part. Participants were asked to find substituted words in texts. The dependent variables were percentage of detected substitution words (accuracy), reading speed, adjusted accuracy (accuracy/reading speed), and preference measures (perception of text legibility, perceived of text sharpness, perception of difficulty in reading, and general preference). There was no significant effect of font type on accuracy, reading speed or adjusted accuracy. On the preference measures, there were no significant effects of font type on text legibility and perceived text sharpness. There was an effect on perception of difficulty in reading and general preference, Times New Roman was perceived as more difficult to read than Arial. On general preference, at the same font size and text style, Arial was more preferred than Time New Roman.

Ling and Van Schaik (2006) conducted two experiments to the effect of font type and other variables on visual search and information retrieval respectively. The independent variables included two levels of font types Arial (10 point) and Times New Roman (12 point) (the difference in point size in order to have the same physical size of text) presented on a Intel Pentium, 333 MHz, 64MbRAM, Microsoft NT4 operating system, 14 inch monitors. In the first experiment, there were 72 participants, a majority (61%) aged 25 years and under and the remainder aged 26 to 50 years. The participants were asked to find a target hypertext link from five hypertext links on the content area of the web page. The participants were asked to undertake the task as quickly and as accurately as possible. There was no significant effect of font type on the performance measures. However, there was a significant effect on the preference measure: participants preferred Arial significantly more than Times New Roman.

In the second experiment, participants were asked to browse web sites in order to find the answer to a question. Each web site had a hierarchy of three levels. The correct answer for each web site was found from one or two links away from its homepage. There were 99 participants, with a mean age of 24 years. Again, font type had no significant effect on the performance measures. On the preference measure there was a trend (0.10 > p > 0.05) for participants to prefer Arial over Times .

Beymer, Russell, and Orton (2008) used eye tracking to investigate the effect of font type on reading behavior and comprehension from a screen. They compared a Sans Serif font (Helvetica) with a Serif font (Georgia). Strangely, they provide no information about the screen used to present reading material to the participants, although they

mention that they were single page stories. The dependent variables were first pass reading speed ((defined as 1st pass gaze duration / characters read), regression rate, time in return sweeps, fraction of the material re-read, saccade length, and fixation duration. There were 82 participants, Most were younger adults, precise details of age are also not provided. The participants were asked to read a passage for comprehension and to answer multiple choice questions afterwards. There was a trend $(0.10 > p > 0.05)$ that Georgia was read faster than Helvetica but it did not reach significance. There was also no other significant effects of font type.

2.8.4 Previous research on effects of font size in reading from screens

Bernard, Liao, and Mills (2001a, 2001b) conducted a study which investigated the effect of font size (12 and 14 point) presented on a Pentium II based PC computer with a 60 Hz, 96dpi 15 inch monitor. An another independent variable was two levels of font type: serif fonts (Time New Roman and Georgia) and san serif fonts (Arial and Verdana). 27 older participants, age range 62 to 83 years, were asked to find substituted words (see discussion of previous paper by Bernard et al 2001) in the texts. The dependent variables were reading efficiency (the percentage of accurately detected substituted words in the passages, divided by the time taken to read the passages), reading time, and participants' preference. Font size had significant effect on reading efficiency, 14 point size had higher reading efficiency than 12 point size. On reading time, there was a significant interaction effect, 12-point serif fonts were significantly slower to read than the 14-point serif fonts or the 14-point sans serif fonts. On participants' preference, in the same font type, 14-point was significantly preferred than 12-point.

Chadwick-Dias, McNulty and Tullis (2003) conducted two studies comparing the reading behavior of younger (under 55 years) and older adults (aged 55 years and over) that included font size as a variable. The first study included three levels of font size : smallest, medium, and largest (These were three from five text sizes according to function on Internet Explorer 6.0). There were 27 participants in the study, the researchers did not provide specific number of younger and older people but only said that participants were recruited and balanced by age and computer and web experience. In addition, the researchers did not provide specific information about the display screen used, but mentioned that texts were displayed in 800 x 600 resolution on a 17-inch monitor, using Microsoft Internet Explorer version 6.0. Each participant was asked to complete 15 tasks, 5 tasks in each font size condition. The font size was

controlled by the web browser. Older participants took significant longer time to complete the tasks than younger participants. Older participants also had a lower task success rate than younger participants. An overall performance score was calculated from time used and task success. On this measure there was a significant correlation with age. As age increased, the overall performance score decreased. In addition, there was no significant effect of text size on performance for both younger and older participants. The interesting results were that text size did not show a significant effect on participants' performance for either younger or older adults. However, on preference measure, older participants significantly preferred the larger text sizes in comparison to younger participants.

2.8.5 Previous research on effects of text and background colour in reading from screens

Snyder, Decker, Lloyd and Dye (1990) conducted three experiments on the effects of positive and negative contrast on visual search and reading tasks when using a Tektronix GMA201 high resolution monochrome CRT with a 48-cm diagonal screen. The first experiment investigated the effects of positive and negative contrast on visual search task. Positive contrast is light text on a dark background while negative contrast is dark text on a light background. There were 10 participants, although not all the participants took part in the first experiment and it is not clear what the number of participants was or what their characteristics were. The results showed that negative contrast created better performance, ranging from 2 percent to 31.6 percent improvements.

Hill and Scharff (1997) studied the effect of text and background colours on reading from a Macintosh Power PC 7200/120 computer. The independent variables included six levels of text and background colours (yellow on blue, white on blue, red on green, black on grey, black on white, and green on yellow). There were 43 participants in the study, no information is provided about their ages or characteristics. Participants searched for target words in texts ranging in length from 130 to 150 words. There was a significant main effect of text and background colour. Green on yellow provided fastest reading time while red on green provided slowest reading time. In addition, there was a significant interaction between text and background colour and font type. Courier New with green on yellow and Times New Roman with green on yellow created the fastest reading times while Arial with green on yellow, and Courier New with red on green created the slowest reading times.

There was also a significant interaction between text and background colour, font type, and word style (italic or not italic). The researchers concluded in this interaction that there were certain font types that work well with certain word style and background colours.

A further experiment reported by Hill and Scharff (1997) compared black text on different levels of grey background, being black on white, black on light grey, black on medium grey (Netscape default colour), black on dark grey, black on very dark grey, presented on a Macintosh Power PC 7200/120 computers. There were 21 participants, precise details of age are also not provided. The procedure was the same as the experiment discussed in the previous paragraph. Black text on white background was read significantly slower than black on medium grey or black on dark grey. In addition, there was a significant interaction between text and background colour, and font type. Times New Roman black text on medium grey background was read fastest while Times New Roman on very dark grey was read slowest. In addition, there was a significant interaction between text and background colour, and word styles. Italic text on very dark grey background was read slowest while non-italic text on medium grey and italic text on dark grey were read fastest.

Hall and Hanna (2004) investigated four combinations of text and background colour: black on white, white on black, light blue on dark blue and cyan on black. The dependent variables for the academic website were: readability, retention and aesthetics while for the commercial website were: readability, retention, aesthetics and behavioral intention. Results showed significant effects for readability for both types of website, cyan on black produced lower score than other colour combinations. There was no significant different for retention. There was marginally significant effect of aesthetics on the academic web site: light blue on dark blue was given a higher aesthetic score than black on white. There were no other significant differences.

The researchers suggested that for academic websites, on which readability is more of a concern, black on white should be used as it provided a high contrast ratio and participants were more familiar with black on white. For commercial websites, on which aesthetics and intention are more of a concern, coloured text and background combination should be used. Light blue on dark blue was more recommended.

This study provided interesting results about colour for academic and commercial web site. However, a limitation of the study was a small number of combinations of text and background colours investigated and the use of only student participants.

Ling and Van Schaik (2002) investigated the effect of text and background colour on visual search task with personal computers (Intel Pentium, 333 MHz, Microsoft NT4 operating system, 14 in. monitors). Colour combinations investigated were black on white, blue on white, blue on yellow, yellow on blue, red on green, and green on red. The dependent variables were accuracy (number of hits, number of correct rejections), speed (reaction time for hits, reaction time for rejections), and two subjective measures. For the first measure, 15 possible pairs of combinations of text and background colours were presented. Participants were asked to choose which colour combination they preferred. For the second subjective measure, participants judged each combinations of text and background colour on a 9 point rating item. There were 29 participants who were students, mean age of 24 years. The participants searched for a target hypertext link from 10 hypertext links in navigation frames, completing the tasks as quickly and as accurately as possible, but within 5 seconds.

There was significant effect of combination of text and background colour on number of hits (correct response when a target word was presented). Blue on white was less accurate than green on red, blue on yellow, and yellow on blue. There was also a significant effect of combination of text and background colour on reaction time for hits Green on red was slower than other combinations of colour. Yellow on blue was faster than red on green and blue on yellow.

In terms of the number of correct rejections, there was a significant effect of combination of text and background colour, participants performed less accurately with blue on white than other combinations.

On the subjective measures, there were significant differences in preferences. Participants significantly preferred blue on white compared to other combinations of colours, except for black on white. Black on white and black on yellow were more preferred than the other combinations. Green on red was more prefer than red on green. In addition, there was a significant effect of combination of text and background colour on perceived display quality. Blue on white was perceived better display quality than other combinations, except black on white. Black on white and black on yellow were perceived better display quality than the other combinations.

Gradisar, Humar, and Turk (2007) investigated the legibility of text and background colour combinations when reading from a 21" Dell CRT display with screen resolution of 1280 x 1024 pixels. The independent variables were 56 combinations of text and background colours from white, yellow, red, magenta, blue, cyan, green, and black. The results from 468 participants, who were university student, were analysed. The

Participants were divided to six groups and asked to indentified characters from each of six particular groups of combination of colour which sorted by decreasing luminance contrast. Each participant spoke aloud characters which were displayed on screen. The size of characters decreased from 3.8 mm in the first row to 1.4 mm in the last row. The line spacing between rows were varied from 5.8 mm to 3.9 mm. There was one blank space between each character.

The best results were from the combinations of yellow on black, cyan on black, white on blue, black on yellow, white on black, and green on black. The worst results were from combinations of black on blue, red on magenta, green on cyan, and yellow on white. There was a significant effect of colour combinations and effect of polarity in mean score. Overall, there was a significant difference between combinations from the first three groups, which had more luminance contrast, with combinations from the last three groups, which had less contrast. A darker text on lighter background resulted in higher mean numbers of correctly indentified characters than a lighter text on a darker background. When doing an analysis by divided combination of colours to two group: positive polarity (dark text on bright background), and negative polarity (bright text on dark background), the results showed that in the first two groups with had more contrast; the negative polarity significantly had better numbers of correctly indentified characters than the positive polarity. However, in the four groups which had lower contrast, the positive polarity significantly had better numbers of correctly indentified characters than the negative polarity.

When doing an analysis with non-chromatic colour (black or white), Firstly, an analysis only color combinations with either black or white color for the background and one of the remaining seven colors for the text showed that combinations with black background had a significant better numbers of correctly indentified characters than combinations with white background. However, when doing an analysis with the color combinations with either black or white color for the text and one of the remaining seven colors for the background, the results showed that the combinations with black text had a significant better numbers of correctly indentified characters than combinations with white text.

Greco, Stucchi, Azvagno and Marino (2008) conducted three experiments on the effect of text and background colour combinations on legibility, and pleasantness. The first experiment investigated legibility on an Acer Travelmate 803 LCi. The 27 colours in the Microsoft PowerPoint palette were used, 13 colours were assigned to the "dark" category and 14 colours were assigned to the "light" category. Words in Japanese were

used as stimuli to avoid any automatic activation of reading processes. There were 30 participants aged between 18 and 56 years. Participants gave ratings of legibility on a 3 point rating scale (1 = unsatisfactory, 2 = passable, and 3 = excellent). There were 702 stimuli. The mean legibility score was correlated with luminance contrast between text and background, light text on a dark background and dark text on light background had average legibility ratings significantly higher than light text on light background and dark text on dark background. In addition, dark text on light background had the best rating. A further analysis of the dark text on light background found that black and blue were rated the most legible text colours while violet was rated as the least legible text colour. In addition, another analysis of light text on dark backgrounds found that brown, green, blue, and black were rated as providing the most legible background colours while red and violet were rated as the least legible background colours.

The second experiment aimed to investigate participants' rating of pleasantness using the same colour combinations as in the first experiment. However, the authors divided combination of colour to four groups of polarity: dark on dark, light on dark, light on light, dark on light. There were 30 participants aged between 18 - 55 years. Participants rated the stimulus according to their pleasantness on a 3 point rating scale (1 = ugly, 2 = passable, and 3 = very fine). There were significant effects of polarity. All groups of polarity were significantly different from each other. Dark text on a light background and light text on dark background got the best pleasantness rating scores, respectively. A further analysis of dark text on light background found the effect of text colour and the interaction between text colour and background colour. Black and blue were rated as the most pleasant text colours, light red was rated as the most pleasant background colour, while any text colour combination with yellow background was rated as not as pleasant. Another analysis of light text on dark background found an effect of background colour and the interaction between text colour and background colour. Black and blue were the most pleasant background colours.

From these two experiments, Greco et al (2008) showed that the greater contrast between text and background colours, the higher legibility score. Dark texts on light backgrounds provided better legibility. On the pleasantness aspect, the authors found that dark texts on light backgrounds were more pleasant than light texts on dark backgrounds.

This study provides interesting results as they investigated both on legibility and beauty (the author called pleasantness but it was beauty aspects). The participants in the research were a wide range of ages (18 to 55 years (Mean=37.4 SD=11.2), and 18 to

56 years (Mean=29.4 SD=13.5), respectively). However, the age range only just reaches the minimum age for older people in some definitions (people who aged 50 years or 55 years and over).

2.8.6 Conclusions

This review of research on text presentation variables on computer screens and their effects on reading behavior and preferences has shown that while there is a considerable amount of research, it is very thinly spread over the large number of variables, screen types and dependent variables. Much more research is needed with modern computer screens to create a robust corpus of evidence. In particular, very little research has been conducted with older people who may be more affected by variations in presentation variables than younger people.

2.9 Effects of different text presentation variables on reading from screen in other languages and writing systems

As presented in Section 2.6, web design guidelines for older people have been suggested for the English language which is written in the Latin alphabet. In addition, most of the research reported in Section 2.8 above, on the effects of a range of text presentation variables on reading from a screen, has been done on the English language and therefore the Latin alphabet. It is not completely clear whether these results can be generalised to other languages which also use the Latin alphabet, although this does seem a reasonable generalisation. However, there is no research could be found on the effects of these variables on reading from the screen in other languages that use the Latin alphabet. It may be some research exists, but that it is presented in the relevant languages, and that my searches in English did not undercover this research.

However, there is also the situation of case of languages which use writing systems other then the Latin alphabet. There the case for generalization is much less clear. There are some studies which have investigated reading text in Chinese, see Section 2.8.1 for details. For example, Chan and Lee (2005) included two levels of line spacing (single and double spacing) and found that participants read text on a 15" CRT display significantly faster with double line spacing than single line spacing. Chan, Tsang, and Ng (2014) used three levels of line spacing (single, 1.5, and double spacing) and found that proofreading time on a 17" LCD display increased significantly with wider line spacing. Wang, Sato, Rau, Fujimura, Gao and Asano (2009) investigated the effects of line spacing on reading from mobile phone screens and found that there was no

significant differences in times for reading and visual search tasks. Interestingly, although this is only a small number of studies, with varying tasks and text presentation variables, the results are in line with results from English. Nonetheless, it is important to conduct research with a wider range of variables and tasks, and on other languages with different writing systems. I have also mentioned that Chinese was the ONLY writing system for which I have found research other than English. However, there are lots of other writing systems, including Thai.

Therefore, in addition to further research on reading from the screen in English , we also need research on the effect of text presentation variables on reading form the screen in other languages and writing systems. The Thai language will be investigated in this programme of research as a language written in a different writing system from the Latin alphabet. As discussed in section 2.x, Thailand has an increasing population of web users, and a rapidly increasing number of older people, many of whom will be older web users in the near future. Currently there are no web design guidelines which have been developed for the Thai language or for older Thai web users.

2.10 Conclusions

The literature review reported in this chapter presented the characteristics and demographic of older people, and the different definitions of older people in ICT research. Then the definition of older adults to be used in the current programme of research was developed, based on the method for calculating minimum age for older adults for the UK and Thailand. This chapter also presented the wide range of web design guidelines for older people. After considering and organising the recommendations from current web design guidelines for older people, the recommendations related to the presentation of text on web pages were selected to investigate in a series of empirical studies. These recommendations are also related to the changing visual capabilities in old age, the factor which probably has the greatest impact on how older people use of the web.

Recommendations related to the presentation of text on web pages, specifically line spacing, text justification, font type, font size, and text and background colour, from different web design guidelines were reviewed in specific details. The review found that each web design guidelines usually provides different recommendations. Previous research on each recommendation was also reviewed. There has been little research with older people, in spite of the fact that they are a growing proportion of both the population in general and computer users in particular.

Chapter 3

Effects of line spacing and text justification on reading webpages by younger and older people in Thailand and the UK

3.1 Introduction

This chapter presents the results of the first study in my programme of research, which investigated the effect of line spacing and text justification on reading webpages by younger and older adults both in Thailand and the UK.

Line spacing and text justification were chosen to investigate as independent variables in the first study because both of these aspects of text presentation are mentioned in multiple sets of guidelines, as discussed in the literature review in Chapter 2 (see section 2.7 in particular). However, the recommendations in these guidelines are unclear as to why particular presentations would be better for older adults and very little evidence is available from older adults reading from modern computer screens to support the recommendations.

Line spacing is one of the aspects of text presentation which many web design guidelines usually mention but the recommendations they make about this aspect are varied. There are five sets of guidelines which make recommendations about line spacing:

1) SPRY Foundation (Holt and Komlos-Weimer, 1999) suggests "increasing the white space between two lines of text by even a small amount (1 or 2 points)"

2) Holt (2000) suggests that "older adults may have more trouble reading pages that are single-spaced rather than double-spaced. An alternative is to

format paragraphs at 1½ spaces, or add a few extra points of space between lines"

3) Agelight (2001) suggests that line spacing should be 2 points larger than the typeface

4) The SilverWeb guideline does not give any specific detail, suggesting only that "there should be spacing between the lines" (Zaphiris, Kurniawan and Ghiawadwala, 2007, p. 69; see also Kurniawan and Zaphiris, 2005)

5) The National Institute on Aging/National Library of Medicine (NIA/NLM) (2002) specifically suggests that the line spacing must be double spaced.

Unlike the recommendations on line spacing, the recommendations on text justification are all point in the same direction. Left-justified text is recommended by six sets of guidelines (Agelight, 2001; Holt and Komlos-Weimer, 1999; Kurniawan and Zaphiris, 2005; NIA/NLM, 2002; Zaphiris, Kurniawan and Ghiawadwala, 2007; Zhao, 2001). The other sets of web design guidelines do not mention text justification. In spite of this agreement between sets of guidelines, text justification was selected as an aspect to investigate in the first study because there appears to be no evidence to support this recommendation. Moreover, text justification has usually been investigated together with line spacing in research about reading text on screen (Kolers, Duchnicky, and Ferguson, 1981; Kruk and Muter, 1984). Ling and van Schaik (2007) investigated text justification in relation to reading text on the web, but this study did not include any older participants.

The current study used a range of combinations of line spacing settings and text justifications to present content on a website. The participants were both older and younger people in both the UK and Thailand. The UK participants participated using an English website while the Thai participants participated using the same website, but translated into Thai. This range of participants allowed me to investigate which of the tested combinations is most appropriate for older web users in both Thailand and the UK.

The study also was the first evidence-based research for constructing web design guidelines for Thai language web sites, as currently there are no web design guidelines for Thai websites, whether for younger or older users.

The following research questions were addressed by this study:

1. Does line spacing have an effect on reading performance and preferences of younger and older people when reading webpages?

2. Does text justification have an effect on the reading performance and preferences of younger and older people when reading webpages?

3. Does age group have an effect on reading performance and preferences when reading webpages?

4. Does participants' nationality and their language have an effect on reading performance and preferences when reading webpages?

5. Do attitudes toward the web have an effect on their reading performance when reading webpages?

3.2 Method

3.2.1 Design

A four way mixed design was used in this experiment. Age Group and Nationality were the between participant independent variables, and Line Spacing and Text Justification were the within participant independent variables.

The independent variables had the following levels:

- Age Group - participants were either Older Adults (55 years and over for participants in Thailand and 65 years and over for participants in the UK, see Section 2.3 for calculation of appropriate minimum age for older adults in Thailand and the UK) or Younger Adults (18 to 39 years in both Thailand and the UK)

- Participants nationality, language and writing system (Nationality for short) - participants were either British people in the UK who were native speakers of English or Thai people in Thailand who were native speakers of Thai

- Line Spacing – single Line Spacing (1S), 1.5 Line Spacing (1.5S), or double Line Spacing (2S)

- Text Justification – left-only justified (L) or left-right justified (LR)

Each participant undertook six tasks on a website about the Olympic Games, one task with each of the six combinations of Line Spacing and Text Justification. Tasks were to find particular items of information about the Olympic Games, such as names of successful athletes, sport rules, and records. The order of presentation of the six combinations was counterbalanced between participants to compensate for practice and fatigue effects.

Three dependent variables related to performance were measured:

- Time spent per webpage – this was a measure of the reading speed of users on a website.

- Number of webpages visited - this was a measure of the efficiency

- Percentage of correct answers – this was a measure of the percentage of participants who could complete the task correctly.

Time spent per webpage was a primary dependent variable as the main investigating in this study. The Number of webpages visited and Percentage of correct answers were dependent variables which also measured in order to check for speed-accuracy tradeoffs in the way the participants undertook the tasks.

Four dependent variables related to participants' preferences were measured:

- Participants' attitude towards the web were measured using the Attitudes toward the Web Scale (Burn, 2003) which has three factors: Confidence, Performance, and Fashion (see section 3.2.4.1, below for more detailed information)

- Participants' ratings of each Line Spacing condition on three dimensions: Ease of Reading; Pleasantness of Reading; and Speed of reading

- Participants' ratings of each Text Justification condition on three dimensions: Ease of Reading; Pleasantness of Reading; and Speed of Reading

- Participants' ratings of their overall preference for each of combination of Line Spacing and Text Justification

The preference dependent variables were all measured on five point Likert items.

3.2.2 Participants

60 people participated in this experiment. 24 participants in the UK and 36 participants in Thailand. One Older participant in the UK dropped out of the experiment halfway through the study, so another participant was recruited to replace him.

The UK participants comprised 12 Younger and 12 Older Adults. The UK Younger Adults comprised 9 males and 3 females, aged between 24 and 31 years (Mean = 26.42 years, SD = 2.39). Nine were Ph.D. students and three were employed. The UK Older Adults comprised 8 males and 4 females, aged between 65 and 78 years (Mean = 72.17 years, SD = 4.02). All the UK Older participants were retired.

The Thai participants comprised 18 Younger and 18 Older Adults. The Thai Younger Adults comprised 6 males and 12 females, aged between 20 and 38 years (Mean = 28.22 years, SD = 4.81). Two were undergraduate students, 7 were Masters students, six were Ph.D. students and three were employed. The Thai Older Adults comprised 6 males and 12 females, aged between 60 and 76 years (Mean = 63.67 years, SD = 3.91). Two were employed and the other 16 were retired.

The Younger Adult participants were offered a gift voucher valued at £10. The Older Adult participants were offered a gift voucher valued at £15, as the sessions for the Older Adults took considerably longer than those for Younger Adults.

The UK Younger Adult participants were recruited by sending emails to students in Department of Computer Science at the University of York. I also asked friends who were studying at the University of Manchester to invite their British friends to participate in the study. The UK Older Adult participants were recruited from the list of older people who had participated with previous studies for the Human - Computer Interaction Research Group in the Department of Computer Science at the University of York.

The Thai Younger Adult participants were recruited from the members of the Thai Society at the University of York and students at Suranaree University of Technology in Thailand. The Thai Older Adult participants were recruited from the Association of Retired Staff of Nakhon Ratchasima Rajabhat University in Thailand.

In both the UK and Thailand, once the Older Adult participants had taken part in the study, they were asked to invite their friends if they would like to participate in the study.

3.2.3 Equipment

The study was conducted on a personal computer (Acer Aspire 4741, Intel (R) Core (TM) i5) running Windows 7 Home Premium and Internet Explorer 9 with a standard keyboard and 2-button mouse with a scroll-wheel. The screen size was 1366 x 768 pixels. Morae[1] software was used to record and analyse the sessions.

3.2.4 Materials

3.2.4.1 Pre-study questionnaires

There were two pre-study questionnaires. The first pre-study questionnaire was adapted from the Attitudes toward the Web Scale (Burn, 2003), a set of 18 statements about respondents' attitudes towards the web, which are answered on five point Likert items (1 = strongly disagree, 5 = strongly agree). This scale has three Factors: Confidence, Performance, and Fashion, see Chapter 2, Section 2.4 for more details. The full set of questions and scoring of the scales can be found in Appendix 10. The data collected from this questionnaire provided participants' attitude towards the web, which was one of the preference dependent variables.

The second pre-study questionnaire consisted of a set of questions about the use of the web and demographic information. Questions included information about age, gender, occupation, experience with the web, and use of the web. The questions also asked participants to rate their level of computer experience (on a 7 point Likert items: 1 = none at all, 7= extensive) and their expertise in using the web (on a 7 point Likert items: 1 = none at all, 7=expert). The full set of questions can be found in Appendices 3 and 4, for the English and Thai versions respectively.

3.2.4.2 Post-study questionnaire

A post-study questionnaire measured participants' ratings about Line Spacing, Text Justification and the combinations they had experienced in the study. Participants were

---

[1] http://www.techsmith.com/morae.html

asked to rate how easy, pleasant, and fast it is to read text with the three different Line Spacing levels and the two different Text Justification levels. Participants were also asked to rate the six combinations of Line Spacing and Text Justification to establish which they preferred the most and explain their reasons in an open-ended question. The ratings were all on five point Likert items (1 = least preferred / strongly disagree, 5 = most preferred / strongly agree).

The post-study questionnaire, Appendices 5 and 6, for the English and Thai versions respectively, provided three preference dependent variables: participants' ratings of each Line Spacing on three dimensions, participants' ratings of each Text Justification on three dimensions, and participants' ratings of their overall preference of each of combination of Line Spacing and Text Justification.

3.2.4.3 Olympic Games website

The website content that participants would read in the study needed to be interesting to both younger and older people in the UK and Thailand. However, the content also needed to be sufficiently unfamiliar so that participants could not answer questions based on their prior knowledge. .

Content about the Olympic Games was selected for the website because the Olympic Games is one of the most important international sport competitions involving many events and people from over 200 countries; the study was also planned during the run up to the 2012 Olympic Games in London, so it was very topical at the time. The Olympic Games is divided into two main events: the Summer and Winter Olympic Games. However, most people, especially Thai people, are familiar with the Summer Olympic Games rather than the Winter Olympic Games. In the end, only contents about the Summer Olympic Games were needed for the website.

To create the website, which was named the "North Yorkshire Olympic Initiative", the first source of information was the official website of the Olympic Movement[2]. The second source was a Wikipedia article on the Olympics[3]. After reviewing this source of information about the Olympic Games, a structure of the website was created. The content, approximately 40 A4 pages, was separated into four main sections: *About the Games*, *History of the Games*, *Olympic Sports*, and *Previous Olympic Games*, as shown in Figure 3.1.

---

[2] http://www.olympic.org
[3] http://en.wikipedia.org/wiki/Olympic_Games

Once the content of the website was finalized for the UK study, six versions of the website were created by using Drupal[4], an open source web development software. A separate version of the website was created for each combination of Line Spacing and Text Justification. Figure 3.1. shows an example of a webpage from the website with Single Line Spacing and Left only Justified text. The six versions were all available via the Web. However, the Thai versions of the website were created using Adobe Dreamweaver and were only available on the computer used to run the study, as Internet coverage in Thailand is not extensive. Thus, using offline websites was more suitable for collecting data in Thailand.

In the study, participants read the content on a webpage and selected an appropriate hypertext link for going to the next webpage until they found the answer to the question posed. If each webpage had only one link, participants would know which hypertext link to choose without having to actually read any content on the page. In order to make sure that participants were really reading the text on each webpage and not just looking for a single link, distractor links, links that were not related to the answer and that did not lead to a webpage with the answer, were added. Figure 3.2. shows an example of a correct link (blue rectangle) and distractor links (red rectangle) on a webpage from the website. The websites, both the English and Thai versions, can be found in Appendix 7.

---

[4] http://drupal.org
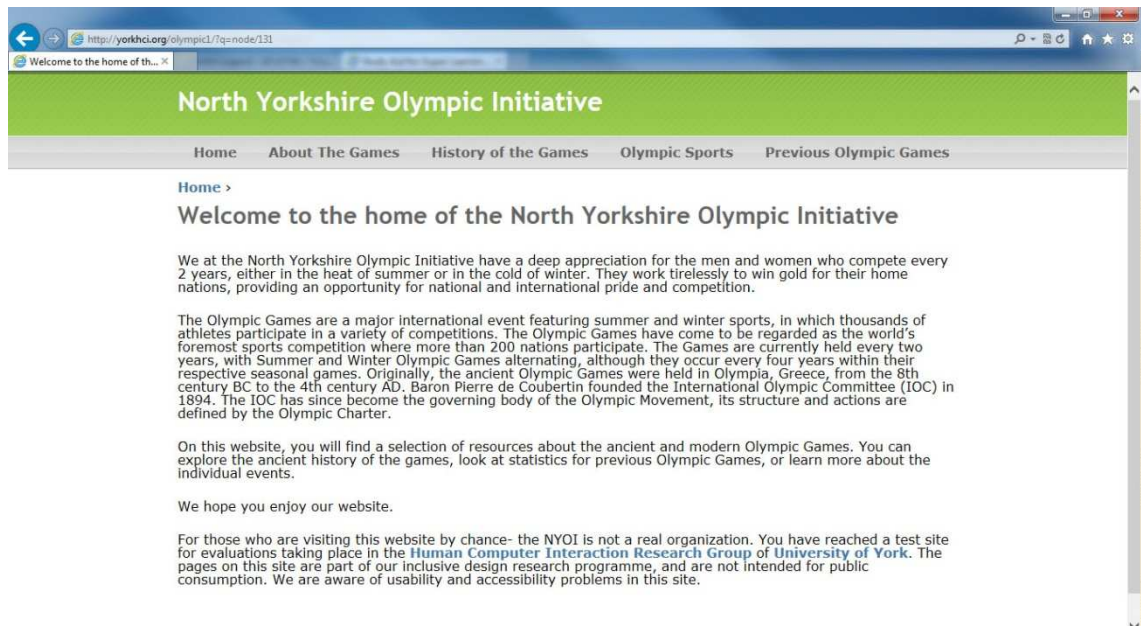
Figure 3.1: Example page from the experimental website (Condition: Single Line Spacing and Left only Justified text)

Figure 3.2: Example of a correct hypertext link (blue rectangle) and distractor links (red rectangles) on a webpage from the experimental website

Figure 3.3 shows that the "About the Games" page had 6 links. These lead to pages about the International Olympic Committee (IOC), Olympic traditions, the Olympic motto, the Olympic rings, Opening ceremonies, and Closing ceremonies. One task asked a question about the Olympic flame for which the answer can be found on the webpage about "The Torches". The link to the webpage about "The Torches" was on

the webpage about "Olympic traditions". The other tasks are described in an Appendices 8 and 9, for the English and Thai versions respectively.

Task 1: How many times has the Olympic flame been carried across the water?

Answer: Twice

Optimal Path: Home > About the Games > Olympic Traditions > Torches

About The Games

| International Olympic Committee (IOC) | Olympic Traditions | Olympic Motto | Olympic Rings | Opening Ceremonies | Closing Ceremonies |

The Torches
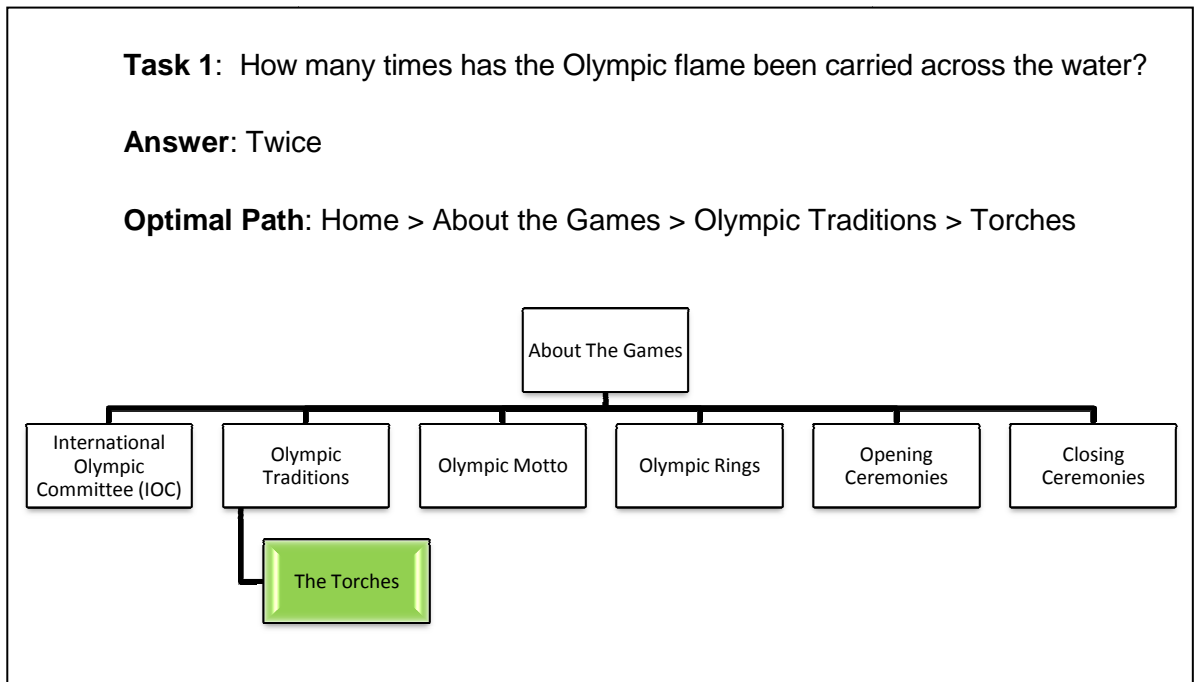
Figure 3.3: Webpages in the section "About the Games" showing the webpage where the answer to Task 1 could be found

3.2.4.4 Experimental tasks

The tasks asked participants to read and find the answers to questions about the Olympic Games. These tasks were divided between the four main sections of the website:

About the Games	1	task
History of the Games	1	task
Previous Olympic Games	2	tasks
Olympic Sports	2	tasks

The section on *Previous Olympic Games* and *Olympic Sports* had more tasks because these contained more content which could be separated to more webpages such as Olympic host countries and sport names. The optimal path to the answer for each task was via three or four webpages (excluding the homepage from which each task started).

3.2.4.5 Translation of materials into Thai

All materials were available in two versions: an English language version for participants in the UK and a Thai language version for participants in Thailand.

For control quality of the translation from English to Thai, the materials were separated into two groups. The first group was data collection materials; for these materials it was critical that the translation was as close to perfect as possible. This was to ensure that data collected in the experiment were comparable between the two groups of participants. In this group were the consent form, the two pre-study questionnaires, the post-study questionnaire, and the texts for the tasks in the study.

The second group was the content on the website. While it was important that the content was translated to a high level of quality, if there were slight differences between the two versions it would be less critical.

I decided to use different techniques for the quality control of the translations. Materials in the first group used back translation for checking the quality of translation. Back translation is usually used for important research documents directly related to data collection such as consent forms, questionnaires, and manuals (Andriesen, 2008).

However, back translation involves a high cost and a considerable amount of time (Andriesen, 2008). Therefore, for materials in the second group I used another method adapted from Mullis, Kelly, and Haley (1996); the committee approach for translation.

Full details of the translation processes and quality controls involved can be found in Appendix 21.

3.3 Procedure

The study was conducted at a number of locations, all quiet rooms at the institutions where the participants studied or where they came to take part in the study.

Before starting the study, participants were briefed about its nature, and their rights. Any questions that participants had about the study were answered. When they were happy about participation in the study, they were asked to sign the consent form section A (see Appendices 1 and 2, for the English and Thai versions respectively). After signing the consent form, participants were asked to complete the pre-study questionnaire (see Appendices 3 and 4, for the English and Thai versions respectively).

Participants were invited to themselves familiar with the computer, monitor, mouse, and the Internet Explorer web browser to be used in the study. When participants were comfortable and ready to start, the researcher provided the first task. After the participant understood the question in the first task, the researcher opened the website with the appropriate combination of Line Spacing and Text Justification and the participant undertook the task. This process was repeated for each task until the participant had completed all six tasks. All tasks were recorded using Morae for later analysis.

The order of presentation of the websites with the six combinations of Line Spacing and Text Justification was counterbalanced between participants.

After completing all the tasks, participants were asked to rate their preference for each combination of Line Spacing and Text Justification using the post-study questionnaire (see Appendices 5 and 6, for the English and Thai versions respectively). As a reminder, examples of all six combinations were provided to participants.

Participants were then debriefed and the purpose of the study was fully explained to them and any questions they had were answered. Participants were then asked to sign Section B of the consent form to show they were happy with their experience.

Each session took approximately 30 minutes to complete for Younger Adult participants, and approximately 45 minutes to complete for Older Adult participants.

3.4 Results

Analysis of Variance (ANOVA) was used to investigate the effects of the independent variables of Line Spacing and Text Justification and the appropriate post-hoc analyses were conducted when there were any significant effects from the overall ANOVA analysis. In addition, for participants' preference ratings, t-tests were used to investigate whether preference ratings were significantly above or below the mid-point of the rating scale.

3.4.1 Time spent per webpage

When conducting the analysis of Time Spent per Webpage, the time which participants spent on the homepage on its initial presentation was not included in the calculation. This was because some participants, especially the Older Adult participants, often spent time for reading the content on the website homepage after they got the tasks, whereas other participants read the content on the homepage only when they got the first task and did not spent much time on homepage after they got the later tasks. Time Spent per Webpage was measured to the closest second.

Firstly, I did the histograms from the data of each combination of Line Spacing and Text Justification. Each histogram showed the distribution of data and provided Mean and Standard Deviation (SD). If the histogram was not a normal distribution, it was necessary to normalise before doing the analysis. For each combination of Line Spacing and Text Justification, any times which were longer than the mean plus two standard deviations (mean+2SD) or shorter than the mean minus two standard deviations (mean–2SD) were adjusted to mean plus/minus two standard deviations respectively. 19 data points out of a total of 360 data points (5.28%) were adjusted in this manner, these were spread evenly across the different combinations.

An four way Analysis of Variance (ANOVA) found that Line Spacing, Text Justification, and Nationality had no significant effect on the Time Spent per Webpage (Line Spacing: $F_{(2,112)}=1.43$, n.s.; Text Justification: $F_{(1, 56)}=0.07$, n.s.; Nationality: $F_{(1, 56)}=2.50$, n.s.). However, there was a significant effect of Age Group ($F_{(1, 56)}=39.94$, $p<.001$, $\eta_p^2 =.42$). There were no significant interactions between any of the independent variables. Older Adults spent significantly longer reading per webpage than the Younger Adults (Mean Younger Adults=14.74 sec. SD=9.04; Mean Older Adults=23.02 sec. SD=7.23). The Younger Adults performed 56.2 per cent faster than the Older Adults.

3.4.2 Number of Webpages Visited

Data on Number of Webpages Visited by each participant in each combination of Line Spacing and Text Justification was normalised by using the same method as described Section 3.4.1, above. 23 data points out of a total of 360 data points on Number of Webpages Visited were adjusted.

For Number of Webpages Visited, there was a significant main effect for Nationality, $(F(1,56)=7.11$, p<.05 η_p^2 =.11), but no significant main effects for Line Spacing, Text Justification or Age Group (Line Spacing: $F(2,112)=0.34$, n.s.; Text Justification: $F(1,56)=1.87$, n.s.; Age Group: $F(1,56)=2.47$, n.s.). There were no significant interactions between any of the independent variables. UK participants visited fewer number of webpages than Thai participants (Mean UK participants=5.90 webpages, SD=2.77; Mean Thai participants=6.63 webpages, SD=3.16).

3.4.3 Percentage of correct answers

For Percentage of correct answers, there were no significant main effects for any of the independent variables (Line Spacing: $F(2,112)=0.91$, n.s.; Text Justification: $F(1,56)=2.50$, n.s.; Age Group: $F(1,56)=2.88$, n.s.; Nationality: $F(1,56)=4.03$, n.s.).

There was a significant interaction between Text Justification and Line Spacing $(F(2,112)=4.07$, p<.05, η_p^2 =.07), and between Text Justification, Line Spacing, and Nationality, $(F(2,112)=4.07$, p<.05, η_p^2 =.07). There were no other significant interaction effects.

Scheffé post hoc analyses were used to test the specific differences in the interactions. Table 3.1 shows the pattern of observed t-values for the two way interaction of Line Spacing and Text Justification combination for Percentage of correct answers. No observed t-value between any combination was greater than the critical t-Scheffe value (95% confidence level: 3.92). So the post-hoc analysis failed to reveal any significant differences in the interaction between Text Justification and Line Spacing, suggesting it was a marginal interaction.

Table 3.1 Observed t - values between all pairs of Text justification and Line Spacing combination for Percentage of correct answers

	1S-L	1.5S-L	2S-L	1S-LR	1.5S-LR	2S-LR
1S-L	-	1.00	a	1.43	a	1.76
1.5S-L		-	-1.00	0.57	-1.00	1.00
2S-L			-	1.43	a	1.76
1S-LR				-	-1.43	0.57
1.5S-LR					-	1.76
2S-LR						-

a. The t cannot be computed because the standard error of the difference is 0.

Table 3.2 Observed t - values between all pairs of Text justification, Line Spacing, Nationality combination for Percentage of correct answers

		UK						TH					
		1S-L	1.5S-L	2S-L	1S-LR	1.5S-LR	2S-LR	1S-L	1.5S-L	2S-L	1S-LR	1.5S-LR	2S-LR
UK	1S-L	-	1.00	a	1.00	a	1.81	a	a	a	a	a	a
	1.5S-L		-	1.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	2S-L			-	1.00	a	1.81	a	a	a	a	a	a
	1S-LR				-	1.00	1.45	1.00	1.00	1.00	1.00	1.00	1.00
	1.5S-LR					-	1.81	a	a	a	a	a	a
	2S-LR						-	1.81	1.81	1.81	1.81	1.81	1.81
TH	1S-L							-	a	a	1.00	a	a
	1.5S-L								-	a	1.00	a	a
	2S-L									-	1.00	a	a
	1S-LR										-	1.00	1.00
	1.5S-LR											-	a
	2S-LR												-

a. The t cannot be computed because the standard error of the difference is 0.

Table 3.2 shows the pattern of observed t-values for the Scheffé post hoc analysis of the three way interaction between Text Justification, Line Spacing, and Nationality Percentage of correct answers. No observed t-values between combinations were significant (95% confidence level: 6.08). So again the post-hoc analysis failed to reveal any significant difference in the interaction between Text Justification, Line Spacing and Nationality, suggesting it was a marginal interaction.

To investigate whether there was a speed-accuracy trade-off in the way the participants undertook the tasks, the Time Spent per Webpage and the Percentage of Correct Answers were correlated. However, this correlation for both Younger and

Older Adults were not significant (Younger Adults: $r(30) = .16$, n.s., Older Adults: $r(30) = -0.20$, n.s.).

3.4.4 Preference measures

The preference measures were participants' ratings of Ease, Pleasantness, and Speed of Reading for each condition of Line Spacing and Text Justification, and their Overall rating of preference of each combination of Line Spacing and Text Justification.

Correlations between the participants' rating on these three dimensions were calculated to investigate whether the ratings were measuring different dimensions of the participants' experience.

Table 3.3 and Table 3.4 show a strong pattern of correlations between all three measures for both the Line Spacing and Text Justification variables. These results meant that participants had only one underlying experience dimension on which to rate the reading tasks.

Therefore a combined User Reading Experience (URE) scores was calculated for each participant for each of the three Line Spacing levels and the two Text Justification levels. This URE scores was the mean of the three ratings for each condition. An added benefit of using this combined score is that scores made up from a number of individual measures are more robust than individual items from participants (Kline, 2000).

Table 3.3 Correlations between ratings of Ease of reading, Pleasantness of Reading, and Speed of Reading for the three levels of Line Spacing

Correlation/Line Spacing	1S	1.5S	2S
Ease-Pleasantness	0.63**	0.67**	0.66**
Ease-Speed	0.68**	0.50**	0.60**
Pleasantness-Speed	0.59**	0.39**	0.54**

** $p<.01$

Table 3.4 Correlations between ratings of Ease of reading, Pleasantness of Reading, and Speed of Reading for the two Text Justification levels

Correlation/Text Justification	L	L-R
Ease-Pleasantness	0.34**	0.66**
Ease-Speed	0.46**	0.46**
Pleasantness-Speed	0.36**	0.50**

** p<.01

3.4.4.1 Analysis of User Reading Experience scores (UREs) for Line Spacing, Age Group, and Nationality

A three way ANOVA on the URE scores for Line Spacing, Age and Nationality found that both Line Spacing and Age group had a significant effect on the URE scores (Line Spacing: $F(2,112)=92.37$, $p<.001$, $\eta_p^2=.62$; Age Group: $F(1,56)=7.38$, $p<.01$, $\eta_p^2=.12$). However there was no significant effect for Nationality, ($F(1,56)=1.18$, n.s.). There were no significant interactions between any of the independent variables.

A Scheffé post hoc analysis was also used to investigate the specific differences between each level of Line Spacing which is illustrated in Figure 3.4. The mean URE scores for single Line Spacing (Mean=2.28 SD=0.72) was significantly lower than the mean URE scores for both 1.5 Line Spacing (Mean=3.84 SD=0.62) (Observed t-value=-14.85, p<.01, critical t at 95% confidence level: 2.48, at 99%: 3.10) and double Line Spacing (Mean=3.97 SD=0.79) (Observed t-value=-10.85, p<.01) while 1.5 Line Spacing and double Line Spacing were not significantly different from each other (Observed t value=-0.99, n.s.).

One sample t-tests against the neutral mid-point score of 3 showed that the mean URE scores for single Line Spacing was significantly lower than neutral (t(59=7.67, p<.001), but the mean URE scores for 1.5 Line Spacing and Double Line Spacing were both significantly higher than neutral (1.5 Line Spacing: t(59)=10.41, p<.001; double Line Spacing: t(59)=9.50, p<.001).
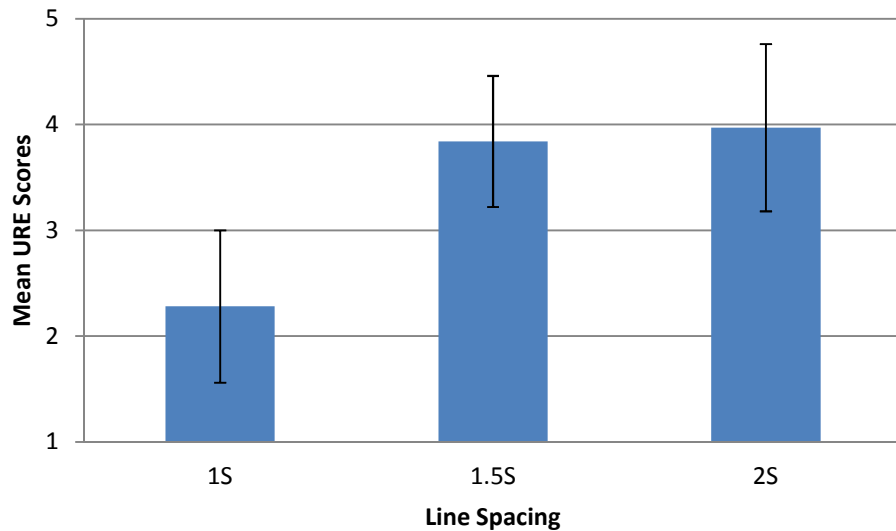
Figure 3.4: Mean URE scores for the three levels of Line Spacing

Overall, Older Adults were significantly more positive in their URE scores than the Younger Adults, (p<.01). The mean URE scores for Older Adults was 3.49 (SD=1.01) and the mean URE scores of Younger Adults was 3.24 (SD=1.07).

One sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for Younger Adults and Older Adults were both significantly higher than neutral (Younger Adults: t(89)=2.10, p<.05; Older Adults: t(89)=4.61, p<.001).

3.4.4.2 Analysis of User Reading Experience scores (UREs) for Text Justification, Age Group, and Nationality

A three way ANOVA on the URE scores for Text Justification, Age and Nationality found that there were no significant differences for Age Group or Nationality (Age Group: F(1,56)=0.18, n.s.; Nationality: F(1,56)=0.52, n.s.). However, Text Justification had a significant main effect, (F(1,56= 6.22, p<.05, η_p^2 =.10). There was also a significant interaction between Text Justification and Nationality, (F(1,56)=13.96, p<.001, η_p^2 =.20). There were no other significant interactions between variables.

Left Justification (Mean Left Justification=3.47, SD=0.57) was significantly lower in URE scores than Left-Right Justification (*p*<.05) (Mean Left-Right Justification=3.79, SD=0.68). One sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for Left and Left-Right Justification were both significantly higher

than neutral (Left Justification: $t(59)=6.38$, $p<.001$; Left-Right Justification: $t(59)=8.99$, $p<.001$).

A Scheffé post hoc analysis was also used to investigate the specific differences in the interaction between Text Justification and Nationality. Table 3.5 shows the observed t-values between each pair of combinations of Text Justification and Nationality. For the Thai participants, URE scores for Left-Right Justification were significantly higher than Left Justification (Mean Left Justification=3.35, SD=0.57; Mean Left-Right Justification =3.98, SD=0.64) (critical t at 95%: 3.47). In addition, Thai participants' URE scores for Left-Right Justification) were significantly higher than the UK participants' URE for Left Justification (Thai mean Left-Right Justification=3.98, SD=0.64; UK mean Left Justification=3.64, SD=0.53). This interaction is shown in Figure 3.5.

Table 3.5: Observed t-values between all pairs of Text Justification and Nationality combination for User Reading Experience scores

		UK		Thai	
		L	**LR**	**L**	**LR**
UK	L	-	0.98	1.96	-3.55*
	LR		-	0.70	-2.74
Thai	L	-	-	-	-4.56*
	LR	-	-	-	-

* p<05
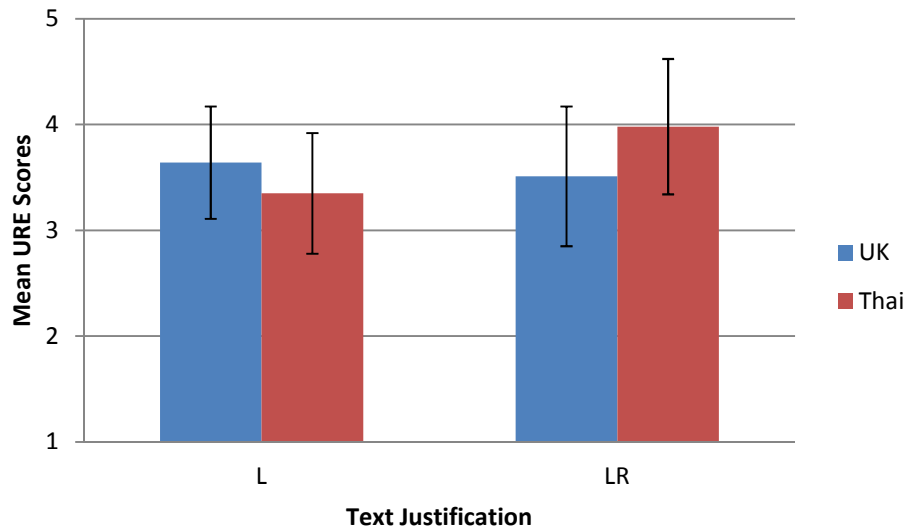
Figure 3.5: Mean URE scores for Text Justification for the UK and Thai participants

One sample t-tests of the URE scores against the neutral mid-point rating of 3 showed that all the mean URE scores were significantly higher than neutral (URE scores on Left Justification for UK participants: $t(23)=5.92$, $p<.001$; URE scores on Left-Right Justification for UK participants: $t(23)=3.81$, $p<.005$); URE scores on Left Justification for Thai participants: $t(35)=3.72$, $p<.005$; URE scores on Left-Right Justification for Thai participants: $t(35)=9.16$, $p<.001$).

3.4.4.3 Participants' overall preference ratings of combinations of Line Spacing and Text Justification

A four way ANOVA on overall preference ratings on the combination of Line Spacing and Text Justification found that Line Spacing, Text Justification, and Age Group all had significant main effects (Line Spacing: $F(2,112)=119.95$, $p<.001$, $\eta_p^2=.68$; Text Justification: $F(1,56)=10.82$, $p<.01$, $\eta_p^2=.16$; Age Group: $F(1,56)=7.42$, $p<.01$, $\eta_p^2=.12$) while Nationality did not have a significant effect, $(F(1,56)=0.59$, n.s.$)$. There was a significant interaction between Line Spacing and Nationality, $(F(2,112)=3.10$, $p<.05)$. There were no other significant interaction effects.

A Scheffé post hoc analysis showed that overall preference ratings for single Line Spacing (Mean rating=1.99, SD=0.94) were significantly lower than both 1.5 Line Spacing (Mean rating=3.79 SD=0.92) (Observed t value=-15.00, $p<.01$, critical t at 95%: 2.48, at 99%: 3.10) and double Line Spacing (Mean rating=3.80, SD=0.93)

(Observed t =-11.92, p<.01) while 1.5 Line Spacing and double Line Spacing were not significantly different from each other (Observed t =-0.59, n.s.), This interaction is shown in Figure 3.6.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean rating for single Line Spacing was significantly lower than neutral (t(119)=11.76, p< .001), but the mean ratings for 1.5 Line Spacing and double Line Spacing were both significantly higher than neutral (1.5 Line Spacing: t(119)=9.47, p<.001; double Line Spacing: t(119)=9.41, p<.001).
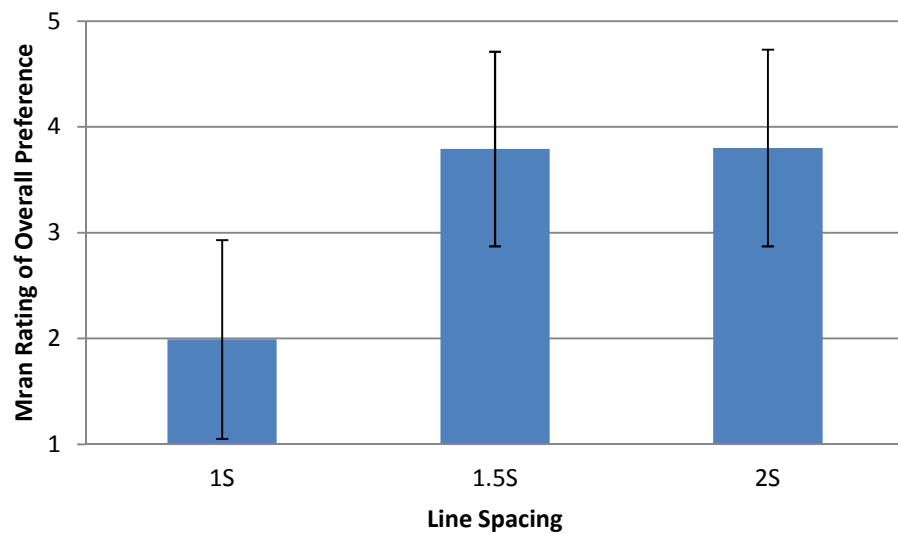
Figure 3.6: Mean Rating of overall preference for three levels of Line Spacing

The main effect for Text Justification was that overall ratings for Left-Right Justification was significantly higher than for Left Justification (Left-Right justification: Mean=3.42, SD=1.25; Left justification: Mean=2.98, SD=1.23).

One sample t-tests against the neutral mid-point rating of 3 showed that the mean rating for Left-Right justification was significantly higher than neutral (t(179)=4.34, p<.001) but the mean rating for Left justification was not significantly higher or lower than neutral (t(179)=0.18, n.s.).

The main effect for Age Group was that Older Adults were significantly more positive in their overall ratings of preference than the Younger Adults (Older Adults: Mean=3.33 SD=1.21; Younger Adults: mean=3.06 SD=1.30).

One sample t-tests against the neutral mid-point rating of 3 showed that the mean rating for Older Adults was significantly higher than neutral (t(179)=3.64, p<.001) but

the mean rating for Younger Adults was not significantly higher or lower than neutral (t(179)=0.63, n.s.).

A Scheffé post hoc analysis was also used to test the specific differences in the interaction between Line Spacing and Nationality. Table 3.6 shows the observed t-value for each combination of Line Spacing and Nationality. For UK participants, single Line Spacing (Mean rating=1.88, SD=1.00) was rated significantly lower than both 1.5 Line Spacing (Mean=4.04, SD=0.94) (Observed t - value=-13.00, p<.01, critical t at 95%: 4.30, at 99%: 5.37), and double Line Spacing (Mean=3.71, SD=0.82) (Observed t - value= -10.22, p<.01) while 1.5 Line Spacing and double Line Spacing were not significantly different from each other (Observed t - value=1.88, n.s.). For Thai participants, single Line Spacing (Mean=2.07, SD=0.89) was significantly lower than both 1.5 Line Spacing (Mean=3.63, SD=0.86) (Observed t - value=-14.63, p<.01), and double Line Spacing (Mean=3.86, SD=1.00) (Observed t - value=-11.51, p<.01), while 1.5 Line Spacing and double Line Spacing were not significantly different from each other (Observed t - value=-1.73, n.s.). This interaction is shown in Figure 3.7.

Table 3.6: Observed t-value between all pairs of of Line Spacing and Nationality combination for User Reading Experience scores

		UK			Thai		
		1S	**1.5S**	**2S**	**1S**	**1.5S**	**2S**
UK	1S	-	-13.00**	-10.22**	-1.11	-8.40**	-10.41**
	1.5S		-	1.88	10.68**	2.50	1.57
	2S			-	9.18**	1.85	-0.88
Thai	1S				-	-14.63**	-11.51**
	1.5S					-	-1.73
	2S						-

* p<.05 ** p<.01
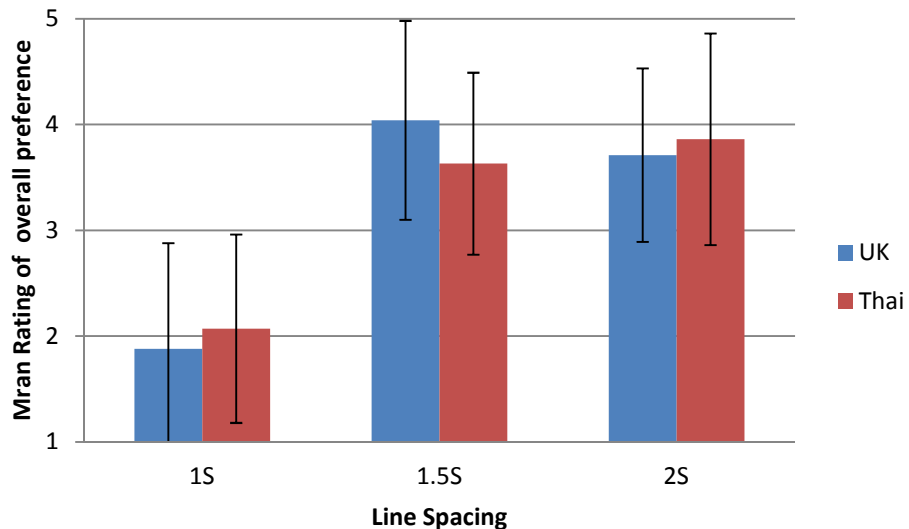
Figure 3.7: Mean Ratings of overall preference for different Line Spacings for UK and Thai participants

One sample t-tests against the neutral mid-point rating of 3 showed that the mean overall preference ratings for single Line Spacing for both UK and Thai participants were both significantly lower than neutral (UK participants: t(47)=7.44, p<.001; Thai participants: t(71)=8.84, p<.001), but the mean ratings for 1.5 Line Spacing and double Line Spacing for both UK and Thai participants were both significantly higher than neutral (1.5 Line Spacing for UK participants: t(47=7.64, p <.001; 1.5 Line Spacing for Thai participants: t(71)=6.15, p<.001; double Line Spacing for UK participants: t(47)= 9.41, p<.001; double Line Spacing for Thai participants: t(71)=7.33, p<.001).

3.4.5 Predicting reading performance from Attitudes to the Web Scale (ATWS)

Some researchers warned that stereotyping people, especially older people, to be a group by age missed the fact that people are so varied in their characteristics (Redish and Chrisnell, 2004). In addition, many researchers suggested a number of interesting dimensions on which older people vary, as discussed in Chapter 2 (Section 2.4). One of these dimensions is attitude towards using the web.

I investigated predicting participants' performance for Line Spacing and Text Justification from their attitudes towards the web. It may be that people who are more positive about the web will perform better in reading from webpages as they think this is a useful and sensible thing to do.

Linear Regression was used to predict participants' reading performance as measured by Time spent per webpage from the three Factors of the Attitudes to the Web Scale (ATWS), Age Group, and Nationality. The Linear Regression used the following formula:

$$\text{Time per Webpage} = a\text{Confidence} + b\text{Performance} + c\text{Fashion} + d\text{Age Group} + e\text{Nationality}$$

The linear regression produced an overall significant prediction ($F(5, 59)=9.05$, $p< .01$, Adjusted $R^2=0.41$). Confidence Factor and Age Group were significant individual predictors (Confidence Factor: $t=-2.46$, $p<.05$) (Age Group: $t=4.41$, $p<.01$). However, Performance Factor, Fashion Factor, and Nationality were not significant predictors. The results are summarized in Table 3.7.

Table 3.7 The B-values, t-values, and significance levels for the linear regression predicting time per webpage from the ATWS factors, Age Group, and Nationality

	B	t	Sig.
Confidence Factor	-3.07	-2.46	.017
Performance Factor	1.71	1.19	.24
Fashion Factor	-0.52	-0.55	.59
Age Group	6.63	4.41	.00
Nationality	-1.73	-1.11	.27

Figure 3.8 shows a scatterplot of the correlation between Confidence Factor and Reading time per webpage. The higher rating on Confidence Factor, the less Reading time per webpage.
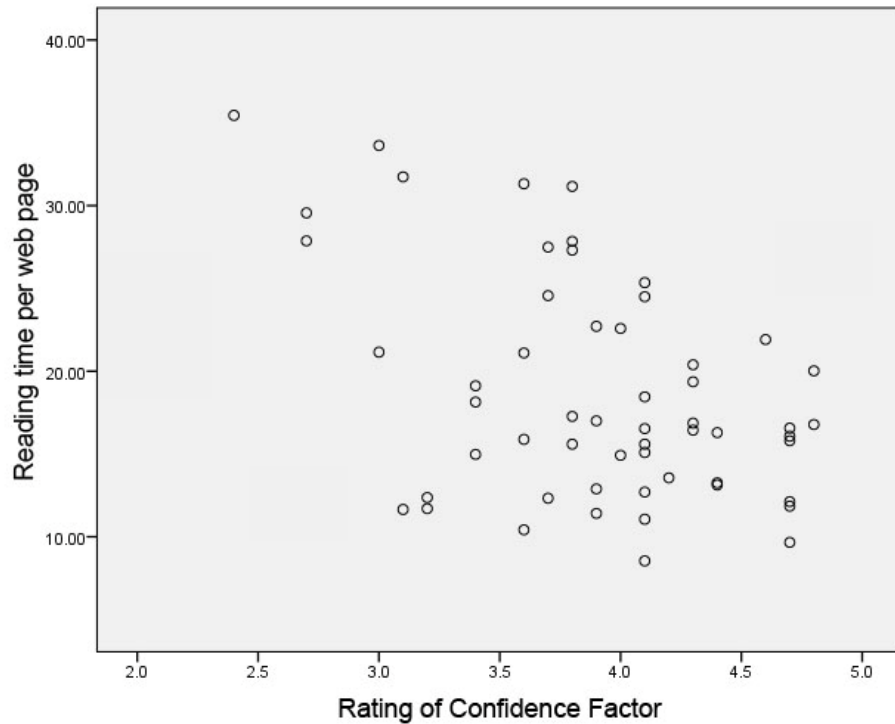
Figure 3.8: A scatterplot of the correlation between Confidence Factor and Reading time per webpage

For the Age Group predictor, Older participants spent longer Reading time per web page than the Younger participants (Mean Older Adults=23.02 sec. SD=7.23; Mean Younger Adults=14.74 sec. SD=9.04). This difference is in line with the findings. .

3.5 Discussion

This study investigated the effect of Line Spacing and Text Justification on performance and preference measures in reading webpages for Younger and Older Adults in the UK and Thailand. The results indicate that there were no effect or a little effect of Line Spacing on the users' performance, but there were differences in preferences. For both UREs, and rating of overall preference, all group of participants preferred 1.5 and double Line Spacing in comparison to single Line spacing, but there was no different between 1.5 and double Line Spacing. In addition, the mean of both UREs and rating of overall preference for single Line Spacing was significantly lower than neutral, while for 1.5 and double Line Spacing were both significantly higher than neutral.

When comparing with previous research about Line Spacing, only one study, conducted by Wang et. al. (2008), investigated effect of line spacing on reading

Chinese text on mobile phone for Chinese older adults. These authors found quite similar findings to the current study, finding no difference on performance measures but significant preference for larger line spacing. However, the authors investigated only two levels of Line Spacing, single and double Line Spacing.

For text justification, there is also no effect of Text Justification in performance measures. On preference measures, for both Thai and the UK participants, Left-Right Justification was significantly preferred in comparison to Left Justification, with Left-Right Justification was rated over neutral but Left Justification was around neutral. However, on UREs, Thai participants significantly preferred Left-Right Justification over Left Justification but there was no difference among UK participants. Moreover, all UREs for both Text Justification for both Thai and the UK participants were higher than neutral.

The reason might be because left - right justification is more familiar to Thai people. Left - right justification not only appears in newspaper, print media and some websites but also when doing electronics or online documents, Thai participants usually organise the text with left-right justification as it is perceived as orderly by them. However, the selection of text justification phenomenon is not distinctly different in the UK.

In terms of generalisability, the participants for the English language website were all English native speakers living in the UK. Whether these results can legitimately be generalised to the presentation of text on websites in languages which use the Latin alphabet other than English is unclear

Turning to the interesting issues about the task in the experiment, firstly, as the researcher who observed participants during the experiment, I noticed that many participants often used a mix of different reading techniques when finding information. Most participants scanned through the links on each webpage and selected a link which they thought would take them to the webpage which had appropriate information to answer the question. In addition, participants would often read quickly through pages that they thought were relevant, skim reading the material. Finally, when participants thought that they had reached the webpage that had the answer for which they were searching, they usually read the text on webpage in a very detailed way. These observations were supported by the data on time spent on task and webpage where users interacted, with some pages being read for very lengthy periods of time whereas some pages were read very quickly, for example, a participant read a webpage 135.96 second but read another webpage only 4.14 second.

These differences in types of reading, and the fact that there were considerable variations in the ways they were used by participants, undoubtedly led to variations in the times spent per page and may have obscured differences due to the independent variables of line spacing and text justification. Essentially the participants were doing a combination of three different tasks, scanning, skim reading and detailed reading, not one of simply "reading" the page.

Secondly, there were issues with some of the questions used in the study. While the tasks were typical for such an informational website, there were multiple places where information might reasonably be found within the small set of webpages. As a result, participants may have spent time on webpages without the answer with the expectation they would find something, thus extending the time they spent there. An example of this phenomenon was a task that asked participants: "How many times has the Olympic flame been carried across the water?". Some participants predicted that the answer should appear on the webpage about "the Olympic torch", which is reasonable, but it was actually on the webpage about "the Olympic flame".

Thirdly during the study I noticed that some participants, particularly older participants, seemed upset when they could not find the appropriate information and took much time for each question. Some participants, particularly older participants, felt lost because they visited many webpages but they could not find the answer or re-visited the same webpage many times. Some participants visited more than 15 webpages in each task while the optimal path was only three or four pages long. These number of webpages were normalized then these data were adjusted. A few participants could not find the webpage which had the answer. They then put the guessed answers which were wrong answers and defeated the purpose of the study. Some older participants gave up in trying to complete the task and dropped out of the experiment.

Finally, some participants said that if they have to find information from a webpage, they do not read all the text on a website but they use the web browser functions such as "find on this page (Ctrl+F)" to look for a specific piece of information. Some participants said they usually use Google in real life for undertaking such tasks. Comments such as these indicate that the tasks used were too artificial for the participants, and likely affected the way participants approached them.

Thus, the next study, Chapter 4, tries to find an appropriate reading task for conducting research into reading on webpages, particularly when working with older participants.

Conclusion: the recommendation on Line Spacing and Text Justification for evidence-based research web design guidelines

Table 3.8 and Table 3.9 show the five criteria which considered for making the recommendation on Line Spacing and Text Justification, respectively. From these two tables, the implications for design of web based text content for web readers in the UK and Thailand are presented below.

Table 3.8: The five criteria which considered for making the recommendation on Line Spacing for Younger and Older Adults in the UK and Thailand

Criteria\Nationality and Age Group	UK		Thai	
	Younger Adults	Older Adults	Younger Adults	Older Adults
Time spent per webpage	no significant difference			
URE	**1.5, 2** > 1			
Testing URE against the mid-point on the 5 point rating scale	1 ↓, **1.5** ↑, **2**↑			
Overall Preference	**1.5, 2** > 1		**1.5, 2** > 1	
Testing Overall Preference against the mid-point on the 5 point rating scale	1 ↓, **1.5** ↑, **2**↑		1 ↓, **1.5** ↑, **2**↑	

A > B means A significantly better than B, = means no significant difference

- means rating is not significantly different from mid-point, ↑ means rating is significantly above mid-point, ↓ means rating is significantly lower than mid-point

* the bold text means it was recommended in that criteria

Table 3.8 shows that all criteria which related to preference measures recommended both 1.5 and double line spacing for Younger and Older Adults in the UK and Thailand.

Table 3.9: The five criteria which considered for making the recommendation on Text Justification for Younger and Older Adults in the UK and Thailand

Criteria\Nationality and Age Group	UK		Thai	
	Younger Adults	Older Adults	Younger Adults	Older Adults
Time spent per webpage	no significant difference			
URE	**Left = Left - Right**		**Left - Right** > Left	
Testing URE against the midpoint on the 5 point rating scale	**Left ↑, Left - Right ↑**		**Left ↑, Left - Right ↑**	
Overall Preference	**Left - Right** > Left			
Testing Overall Preference against the midpoint on the 5 point rating scale	**Left =, Left - Right ↑**			

A > B means A significantly better than B, = means no significant difference

- means rating is not significantly different from mid-point, ↑ means rating is significantly above mid-point, ↓ means rating is significantly lower than mid-point

* the bold text means it was recommended in that criteria

Table 3.9 shows that all four criteria which related to preference measures recommended left - right justification for Thai Younger and Older Adults where as only two criteria recommended left justification. For UK Younger and Older Adults, four criteria recommend left-right justification while three criteria recommended left justification. Thus, recommendation on text justification for UK Younger and Older Adults was both left and lift-right justification.

Proposing 1.5 or double line spacing as a web design guideline did not support some of the previous guidelines as they did not provide specific details (SilverWeb, 2005, 2007) or they recommend other line spacing (SPRY Foundation, 1999; Agelight, 2001; NIA/NLM, 2002). However, this proposing supports the alternative recommendation of Holt (2000) which recommended 1.5 line spacing.

Proposals of left only justification for older adults (Agelight, 2001; Holt and Komlos-Weimer, 1999; Kurniawan and Zaphiris, 2005; NIA/NLM, 2002; Zaphiris, Kurniawan and Ghiawadwala, 2007; Zhao, 2001) are not supported by this study. For all UK web users, both left and left - right justification is recommended.

For all Thai web users, the recommendation on Text Justification is left - right justification, as shown in Table 3.10.

Table 3.10: The recommendation on line spacing and text justification for evidence-based research web design guidelines for younger and older adults in the UK and Thailand

Age group and Nationality	Recommendations on line spacing	Recommendations on text justification
UK younger adults	1.5 or double line spacing	Left and Left - right justification
UK older adults	1.5 or double line spacing	Left and Left - right justification
Thai younger adults	1.5 or double line spacing	Left - right justification
Thai older adults	1.5 or double line spacing	Left - right justification

The results from this study might support the most interesting conclusion that recommendations on line spacing depend on neither the different nationalities and languages nor age group of participants. Furthermore, the recommendation on text justification for UK participants does not play an important role as both text justifications are good for participants. However, the recommendation on text justification is necessary for Thai participants. Left-right justification should be recommended for both Thai younger and older adults.

Chapter 4

Exploring the range and appropriateness of tasks for research about reading webpages

---

4.1 Introduction

This chapter presents the second study in my programme of research, which investigated the range and appropriateness of tasks for research about reading webpages.

In the first study, the participants undertook a reading task that asked them to find specific information on the Olympic Games website to answer questions. The example of the task "How many times has the Olympic flame been carried across the water?". The answer of the question was "twice", it was found on the webpage about "the Olympic flame". This reading task provided interesting results which were presented in Chapter 3. However, there were some problematic issues about this reading task.

As a researcher who observed participants during the experiment, I found that participants used mix of different reading techniques when finding a piece of information. The participants were doing a combination of three different tasks, scanning, skim reading and detailed reading, not one of simply "reading" the webpage. Moreover, there were issues with some of the questions used in the study. There were multiple places where information could be found within the small set of webpages. As a result, participants may have spent time on webpages without the answer with the expectation they would find something, thus extending the time they spent there. In addition, the issues with some questions meant some participants, particularly the older participants, were unhappy and upset during the task. An older participant gave up and dropped out of the experiment. Finally, some participants mentioned that the task did not reflect the way they would usually find the information from the web.

I therefore decided to investigate the various reading tasks of scanning, skimming and detailed reading in more detail. As an initial exercise, a more detailed literature review was undertaken about types of reading on the web. This would help me identify more appropriate tasks to use in the further studies in my programme of research.

4.2 Research on different types of reading on the web

This supplementary literature review focussed on empirical studies about how people read on the web in order to identify the range of tasks that participants undertake in these studies as well as the measures that are typically taken to measure reading behaviour on the web.

4.2.1 Visual search as a reading task

Ling and van Schaik (2002, 2004, 2007), van Schaik and Ling (2001), Pearson and van Schaik (2003), all use a *visual search* task in which participants search for target link words amongst other links in navigation area (Ling and van Schaik, 2002; van Schaik and Ling, 2001; Pearson and van Schaik, 2003) or amongst texts on webpage (Ling and van Schaik, 2004, 2007), and respond by pressing an appropriate key as quickly as possible to indicate that target link word is present or absent from the navigation area or the text on webpages. This task measured on both objective measures (speed and accuracy) and subjective measures (e.g. participants' preference, rating on aesthetic, and quality of screen layout).

Dillon et. al. (2006) also defined their operational task as *visual search* but their visual search task differed from that used by van Schaik and colleagues. In this study, participants were given questions (e.g. "How many books are acceptable for an age 4-7 reading level and cost $21.00") and asked to scan through a set of spreadsheets on the web for counting the number of time which the target information appeared. Thus, there were the two main differences from van Schaik and collegues research: there were more than a single target information on a webpage, and the target information were presented on spreadsheet on a webpage. This task measured on both objective measures (speed and accuracy), and subjective measures (participants' fatigue).

Another type of visual search as a reading task was searching information from a website (van Schaik and Ling, 2003; Ling and van Schaik, 2006; van Schaik and Ling, 2006). In this task, participants were given the questions and asked to visit a website for finding the specific information. Thus, participants visited many webpages in a

website before answering each question. Thus, the current task differs from visual serch task which participants visited only a webpage. This task measured on both objective measures (speed, accuracy, and numbers of links visited), and subjective measures (rating on display quality, aesthetic value, and participants' fatigue).

4.2.2 Skim reading

In some studies (e.g. Muter and Maurutto, 1991), participants were asked to read the text at faster than normal speed in order to get a sense of the content or main ideas of the text. After this task of *skim reading* the text, participants answered multiple choice questions or fill in the blank questions.

In general, skim reading means reading at a much faster than normal speed but there are different operational definitions of skim reading. Muter and Maurutto (1991) defined skim reading as reading at a rate three to four times faster than normal speed. Muter and Maurutto controlled the speed of reading by asking participants to read an instruction sheet while their time was monitored. The participants were then asked to skim the same sheet in one-quarter of the time orginal reading time. After this procedure, the researchers indicated that participants understand the task However, Dyson and Haselgrove (2000) defined skim reading as read at only twice the normal reading speed. Dyson and Haselgrove asked participants to read an initial document at their normal speed while their time was recorded. The participants were then asked to read the next document at twice of the normal speed. If they can speed up their reading, they asked to continue to complete the study. If not, they were asked to read another document with twice of their normal speed. The task measured on only objective measures (speed, level of comphrehension, effective rate (words/min)).

4.2.3 Detailed reading

A detailed reading task was used by Shaikh (2005, see also Shaikh and Chaparro, 2005), which they called reading for comprehension. In these two studies, participants were asked to read the passage as quickly and accurately as possible for comprehension. Boyarski et al (1998) explained that participants read at their usual reading speed this task. After reading, participants then answered multiple choice questions after reading. This task measured on both objective measures (speed, accuracy, and effective reading (score/time)) and subjective measures (participants' preference, participants' rating on ease of reading, sharpness, and ligibility).

Wilkinson and Robinshaw (1987) asked participant to undertake a proofreading task, to identify a particular kind of error in the text by speaking into microphone. The errors were missing or additional spaces (e.g. 'tomorrowthey', 'tomor row'), double or triple reversions (e.g. 'tomrorow', 'toromrow') one missing or one additional letter (e.g. 'tomrrow', 'tomorroww'), misfits (e.g. 'tom #rrow', 't*ey'), or an inappropriate or missing capital, (e.g. 'tOmorrow', 'london'). This task measured on objective measures (speed, and percentage of error missed).

Gujar, Harrison, and Fishkin (1998) modified the proofreading task by asking participants to detect substitution words in the text which varied grammatically but rhymed with the original word, (e.g. "cake" could be replaced with "fake"). The task of detecting substitution words was also adopted for use by many researchers (e.g. Bernard, Liao, and Mill, 2001; Bernard, Fernandez, and Hull, 2002; Bernard et al., 2003). This task measured on objective measures (speed, reading distance, and error rate) and subjective measures (participants' rating on ease of reading)

Dillon, Richardson and McKnight (1990) and O'Hara and Sellen (1997) used text summarisation tasks. Participants were asked to read the text with no time constraints. Participant could re-read the text as often they wished until they were satisfied that they had understood the text. They were then asked to summarise the main points of the text. This task measured on objective measures (reading time and comprehension socres) and subjective measures (participants' opinions on issue such as the range of facilities offerred, improvement they would like, problems they encountered)

4.2.4 Taxonomy of Reading Tasks

In each of the studies reviewed above, the researchers generally used one or at most two different reading tasks. Surprisingly, very little has been written or investiaged about these different types of reading task and how they relate to each other and what underlying concepts they relate to. Dillon (2004) reviewed empirical research on reading from screen in comparison to reading from paper and found that the most of research reflected two implicit views of the typical reader. The first view was of a reader who scanned short texts, searching for typographical errors or other mistakes. The second view was of a reader who was an explorer who searched for a target from information which presented on the screen or on paper. However, Dillon (2004) argued that these two views alone were not sufficient to explain the reality and totality of reading situation.

From the research reviewed, a taxonomy of reading types emerges that captures the different types of tasks that have been used. In doing this literature review, I found that each researcher had used different reading types in their study and claimed that their selected reading type was appropriate to use in conducting studies about reading text on screen. However, I found that the types of reading used in research about reading text from screen can be broadly categorised into the following types of reading:

Within the literature, a variety of studies used tasks that can be categorised as *scanning*. Scanning is where the reader is searching the text for a particular piece of information without reading the whole text. This type of reading does not necessarily require readers to understand any information in the text in relation to the target information or to answer questions about the information found. The studies on visual search and searching for information are examples of scaning.

Skim reading involves the reader reading at a higher rate of speed than normal. The distinguishing feature of this type of reading from scanning is that the individual is not only trying to find a piece of information, but also understand something of the text, the context and general overview of how the information relates to the rest of the text. Readers who are skim reading are often asked comprehension questions regarding the contents, something that is not done in scanning tasks.

In *detailed reading*, readers read all of the content by being given a variety of detail oriented tasks. These tasks require full understanding the text at the word or sentence level in order to undertake some kind of editing, comprehension or summarisation tasks. Reading for comprehension, proofreading, detecting substitution words, and text summarisation are all examples of detailed reading tasks.

Taking this taxonomy of reading types, the study presented in this chapter will explore the range and appropriateness of a range of reading tasks for use in research about reading on webpages. The key outcome of this study will be to identify which tasks best capture the overall reading experience of participants, and the relationship between the tasks, and then use that information to design tasks for use in future experiments in this programme of research.

4.3 Aims

The study presented in this chapter aims to explore the range and appropriateness of a range of reading tasks for research about reading on webpages which will be used in further research about web design guidelines for older people. Three reading types, scanning, skimming, and detailed reading, will be investigated as a within participant

variable. While line spacing is investigated as a between participant variable. Line spacing is selected beacuse it is one of the aspects of text presentation which many web design guidelines usually mention but the recommendations they make about this aspect are varied, see Section 3.1 for details. Line spacing is a veriable which found the significant effect on participant's performance in previous research (Ling and van Schaik, 2007)

4.4 Method

4.4.1 Experimental Design

The study had a three way mixed design. Line Spacing was the between participants independent variable. Reading Type and Tasks were within participant independent variable. The independent variables had the following levels:

- Line Spacing - Single Line Spacing, 1.5 Line Spacing, or Double Line Spacing
- Reading Type - Scanning, Skimming, or Detailed Reading
- Text - text 1 or text 2

Each participant undertook six tasks, two tasks with each reading type. Order of reading type and the text were counterbalanced between participants to avoid fatigue and practice effects.

Three dependent variables related to reading performance were measured:

- Time spent per webpage – this is a measure of the reading speed of users on a website.

- Percentage of correct answers – this is calculated as a measure of the percentage of participants found the correct target word in Scanning and answer the multiple choice questions correctly in Skimming and Detailed Reading. As there was different number of correct answers for each task, the number of correct answers were calcuated to percentage of correct answers.

One dependent variable related to participants' fatigue was measured:

- Participants' ratings of visual and physical fatigue - adapted from Tyrrell and Leibowitz (1990) and Dillion, Kleinman, Ok Choi, and Bias (2006). This dependent variable was considered as the current study aims to find the range

and appropriateness reading task for further studie which included both younger and older people.

- Participants' perceptions of their use of the reading types

4.4.2 Participants

There were 46 participants, 35 males and 11 females. The mean of age was 24.02 years (SD = 4.8, range 18 - 38 years). 20 participants were undergraduate students, 8 were Masters students, 8 were Ph.D. students and 10 were from outside the university.

For the between participant Line Spacing variable, 16 participants participated in the Single Line Spacing condition, 15 participated in 1.5 Line Spacing condition, and another 15 participated in Double Line Spacing condition.

The participants were offered a gift voucher valued at £10 for their participation.

4.4.3 Equipment and Materials

4.4.3.1 Equipment

Each participant undertook the experiment on a personal computer running Windows XP and Internet Explorer 9 with 21.5 inch LCD Monitor, a standard keyboard, and 2-button mouse with a scroll-wheel. The experiment sessions were recorded by TechSmith's Morae software[1] for later viewing and analysis.

4.4.3.2 Materials

4.4.3.2.1 Website, questions, and target words

A website with three pages was developed to provide a practice task for participants and a further six pages for the main experiment tasks. Each page had 275 words of text about the Olympic Games, separated into four paragraphs. Each text had a target word for the scanning task and a set of four multiple choice questions for the skimming and detail reading tasks. The webpages, target words, and questions can be found in Appendix 11.

---

[1] http://www.techsmith.com/morae.html

The texts had approximately the same ease of reading, the mean Flesch-Kincaid Reading Ease Score[2] was 43.7 (SD=2.7, range 40.4–47.7) and the mean Gunning-Fog Score[3] was 13.72 (SD=1.1, range 12.3–14.9).

All the multiple choice questions was tested for accuracy and difficulty by asking three native English speakers to read the text and answer the questions while referring back to the related text. They also rated each question for difficulty on a 9 point Likert items (adopted from Dyson and Haselgrove, 2000). The results from accuracy and difficulty test show that all testers answered all questions correctly. This showed that every question could be correctly answered by the information in the text. The mean difficulty of question was 2.64 (SD=0.5, range 1.67–3.67) while the mean of difficulty of each set of questions was 2.64 (SD=0.2, range 2.25–2.92).

All target words were nouns which were approximatly the same length in number of characters. The mean number of characters was 9.2 (SD=1.1, range 8-10). The target word appeared once in the text on the webpage on either the left or right of the paragraph in either the third or fourth paragraph of the text.

Three versions of the website were created, one for each of the three levels of Line Spacing: one with Single Line Spacing, one with 1.5 Line Spacing, and one with Double line spacing.

4.4.3.2.2 Visual and physical fatigue questionnaire

The visual and physical fatigue questionnaire was adapted from Tyrrell and Leibowitz (1990) and Dillion, Kleinman, Ok Choi, and Bias (2006). The participants were asked to rate their visual and physical fatigue in five dimensions (question 1-5) and the last question (question 6) asked about overall fatigue. The questions in visual and physical fatigue questionaire are presented in Appendix 13.

The rating were all on a 7 point Likert items (1=strongly disagree; 7=strongly agree).

4.4.3.2.3 Participants' perceptions of their use of the reading types questionnaire

Participants' perceptions of their use of the reading types questionnaire had two short questions. The first question was about the frequency which they perceive that they

---

[2] In education, Flesch-Kincaid Reading Ease Score designed to indicate how difficult a reading passage in English is to understand
[3] In linguistics, the Gunning fog score measures the readability of English writing. The index estimates the year of formal education needed to understand the text on the first reading

use each reading task type (scanning, skimming, and detailed reading) . Another question was about the percentage of time which they perceive that they spend using each reading task type when reading on the web. The total percentage for each question was 100 per cent, See Appendix 20.

4.4.4 Procedure

Before commencing the experiment, participants were briefed about the study and the procedures. They were asked to complete an informed consent form, (see Appendices 1 and 2, for the English and Thai versions respectively), and a demographic questionnaire (see Appendices 3 and 4, for the English and Thai versions respectively).

They were then asked to make themselves familiar with the computer, monitor, mouse, and web browser. The participants were then given a practice task. In the practice task, participants were asked to read the text on webpage by using the three different reading types; scanning, skimming, and detailed reading.

For the scanning practice task participants were given a target word and they were then asked to scan through the text on the webpage. When the participants found the target word, they were asked to press the space bar, say the target word aloud and indicate where the target word was on the page..

For the skimming practice task participants were asked to read the text as quickly and accurately as possible in order to understand the main ideas of the text. They were asked to press the space bar to indicate that they had finished skim reading the text. The participants were then asked to complete four multiple choice questions.

For the detailed reading practice tasks participants were asked to read carefully through the text in order to completely understand the text. They were asked to press the space bar to indicate that they had finished detailed reading. The participants were then asked to complete four multiple choice questions and then write 3 - 4 sentences about the text.

After completing the practice tasks participants were asked to describe the difference between scanning, skimming, and detailed reading, in order to confirm that they understood the difference as I anticipated between these three reading tasks.

Then, participants were asked to do the test series of tasks of scanning, skimming and detailed reading. Each participant experienced only one line spacing condition, with all

tasks being undertaken with the same line spacing. Each participant undertook six tasks, two tasks per each of reading type. After completing each of these pairs participants were asked to rate their fatigue on visual and physical fatigue questionnaires (see Appendix 13). The order of reading type and text was counterbalanced to avoid practice and fatigue effects between the participant groups.

When the participants had completed all tasks, they were asked to complete a participants' perceptions of their use of the reading types questionnaire (see Appendices 20).

Each experimental session took approximately 45 - 60 minutes to complete.

4.4.5 Data preparation

Firstly, I did the histograms from the Data on Time spent per webpage in each of Reading type. Each histogram showed the distribution of data and provided Mean and Standard Deviation (SD). If the histogram was not a normal distribution, it was necessary to normalise before doing the analysis by using the method outlined in Section 3.4.1. In this process, 8 of the total of 346 data points (2.31%) were adjusted.

4.5 Results

The effects of Reading Type and Line Spacing on the Time Spent per Webpage, the Percentage of Correct Answers, and participants' ratings of their Visual and Physical Fatigue were investigated. Analyses of Variance (ANOVA) were used to investigate these effects and Scheffé post-hoc analyses were conducted when there were any significant effects in the omnibus ANOVA analysis which needed further investigation. In addition, for participants' ratings of Visual and Physical Fatigue, t-tests were used to investigate whether these ratings were significantly above or below the mid-point of the rating scale.

4.5.1 Time spent per webpage

An Analysis of Variance (ANOVA) found that the main effects for Line Spacing and Text were not significant (Line Spacing: $(F_{(2,43)}=0.35$, n.s.; Text: $F_{(1,43)}=1.99$, n.s). However, there was a significant main effect for Reading Type, $(F_{(2,86)}=181.63$, p <0.001, $\eta_p^2 = 0.81$). There were no significant interactions between variables.

A Scheffé post hoc analysis was used to investigate the specific differences between Reading Types. A Scheffé post hoc found that each Reading Type was significantly different from each other (critical t at 95% confidence=2.49; at 99% confidence= 3.12). Participants significantly spent shorter reading time in Scanning (Mean=22.76 SD=12.14) than both Skimming (Mean=48.43 S.D.=16.23, Observedt=-11.44, p<.01) and Detailed Reading (Mean=84.30 SD=27.09, Observed t=-15.05, p<.01). Participants significantly spent shorter reading time in Skimming (Mean=48.43 SD=16.23) than Detailed Reading (Mean=84.30 SD=27.09, Observed t=-12.06, p<.01). Figure 4.1 shows the mean times spent per webpage (seconds) for Scanning, Skimming, and Detailed Reading.
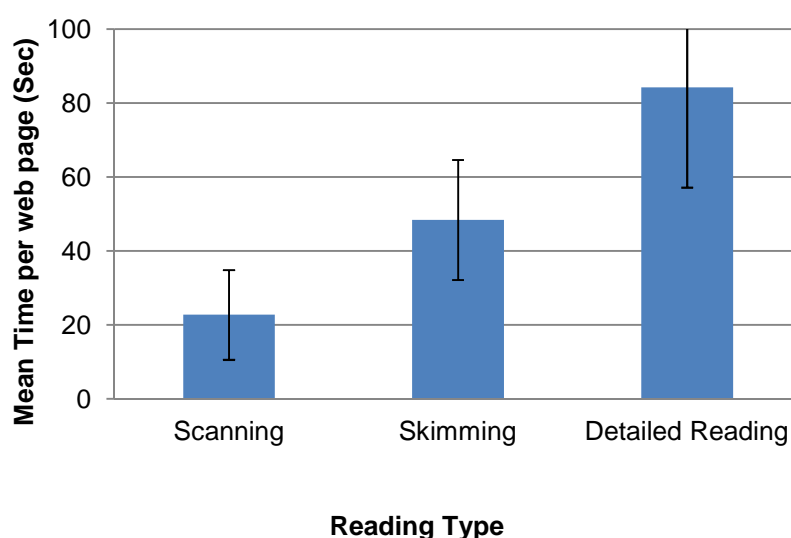
Figure 4.1: Mean Reading Time per webpage (seconds) for Scanning, Skimming, and Detailed reading

4.5.2 Percentage of correct answers

An ANOVA found that the main effect for Line Spacing and Text on percentage of correct answers were not significant (Line Spacing: $F_{(2,43)}=0.11$, n.s.; Text: $F_{(1,43)}=0.35$, n.s). However, there was a significant main effect for Reading Type, ($F_{(2,86)}=106.33$, p<.001, η_p^2 =.71). There was no significant interaction between the variables.

A Scheffé post hoc found that the Reading Types were significantly different from each other. Participants had significantly higher accuracy on Scanning (Mean =100 S.D.= 0) than both Skimming (Mean=49.18 SD=26.33, Observed t=16.97, p<.01) and Detailed

Reading (Mean=68.21 S.D.=25.17, Observed t=10.25, p<.01). While participants had significantly lower accuracy on Skimming (Mean=49.18 SD=26.33) than Detailed Reading (Mean=68.21 S.D.= 25.17, Observed t=-4.49, p<.01). Figure 4.2 shows the mean percentage of correct answers for Scanning, Skimming, and Detailed reading.
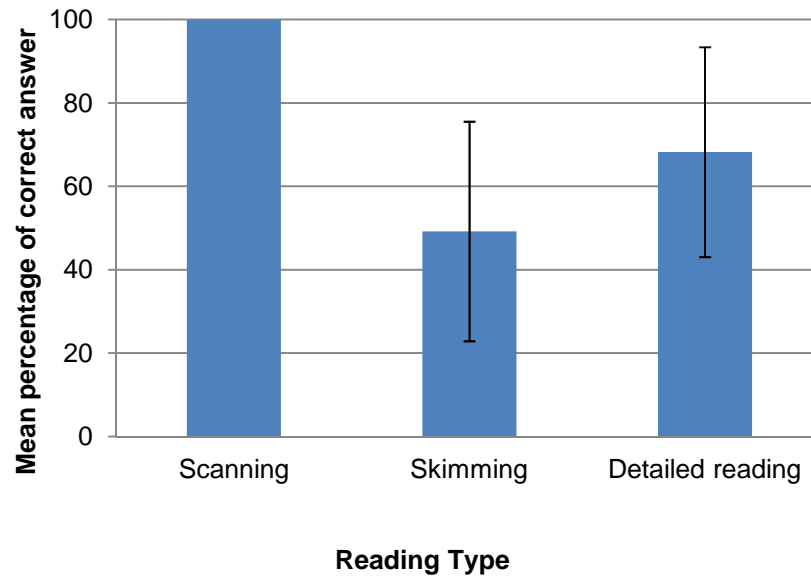
Figure 4.2: Mean Percentage of correct answers for Scanning, Skimming, and Detailed reading

4.5.3   Visual and physical fatigue

Visual and Physical Fatigue were measured the levels of visual and physical fatigue in five aspects (Questions 1 - 5) and the last question for the overall fatigue on a 7 point Likert items. Correlation analyses between the participants' rating on these five aspects (Questions 1 - 5) and overall fatigue (Question 6) were assessed to investigate whether the participants were reacting to different things.

Table 4.1 Correlation between participants' rating on five aspects of Visual and Physical Fatigue and Overall Fatigue for Scanning, Skimming, and Detailed reading

Correlation/Reading task	Scanning	Skimming	Detailed reading
Question 1/Question 6	.56**	.61**	.66**
Question 2/Question 6	.70**	.74**	.69**
Question 3/Question 6	.66**	.70**	.66**
Question 4/Question 6	.83**	.88**	.84**
Question 5/Question 6	.73**	.83**	.78**

* p<.05, ** p<.01

This strong pattern of correlations shown in Table 4.1 suggests that participants' ratings of visual and physical fatigue in the five specific questions were not different from ratings of overall fatigue as measured by Question 6. Therefore the overall fatigue ratings were used in further analyses.

An ANOVA on the Overall Fatigue ratings found that the main effect for Line Spacing was not significant ($F(2,43)=0.27$, n.s.). There was a significant main effect for Reading Type ($F(3,86)=20.90$, $p<.001$, $\eta_p^2 =.33$). There was no significant interaction between variables.

A Scheffé post hoc found that Overall Fatigue ratings for each Reading Type were significantly different from each other. Detailed reading (Mean=4.33 S.D.=1.55) significantly produced a higher Overall Fatigue rating than either Skimming (Mean=3.46 S.D.=1.70, Observed t=-4.05, $p<.01$) and Scanning (Mean=3.07 S.D.=1.50, Observed t=-6.38, $p<.01$). There was no significant difference between Scanning and Skimming (Mean Scanning=3.07 S.D.=1.50; Observed t =-1.91, n.s.). Figure 4.3 shows the Overall Fatigue ratings for Scanning, Skimming, and Detailed reading.

One sample t-tests against the neutral mid-point rating of 4 showed that the mean Overall Fatigue ratings for Scanning and Skimming were significantly lower than neutral (Scanning: t 45) =-4.24, $p<.001$; Skimming: t(45)=-2.17, $p<.05$). While the mean rating for Detailed reading was not significantly different neutral (t (45)=1.43, n.s.).
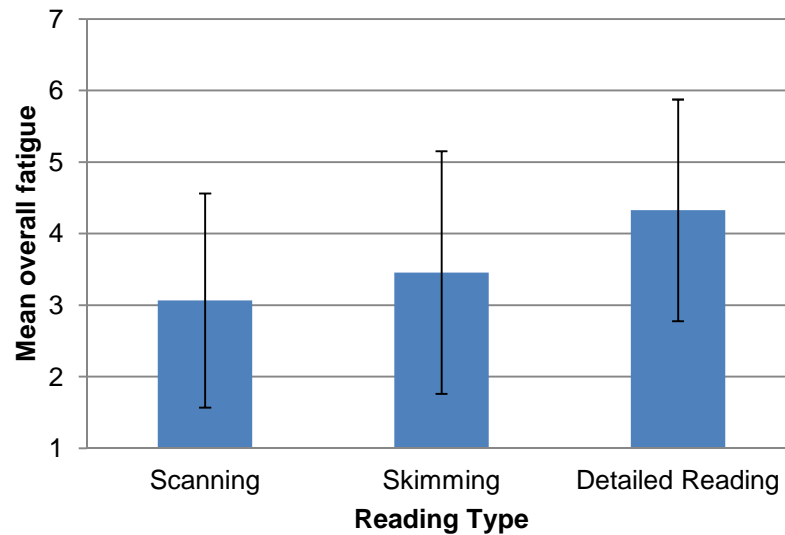
Figure 4.3: Mean Overall Fatigue ratings for Scanning, Skimming, and Detailed reading

4.5.4 Participants' perceptions of their use of the reading types

For the Participants' perceptions of their use of each reading type, one participant completed the questionnaire incorrectly, so their data were omitted from the following analysis. An ANOVA found that the main effect for Reading Type was significant $(F_{(2,88)}=8.19, p<.01, \eta_p^2 =.16)$. A Scheffé post hoc found that participants used Skimming (Mean=43.69 S.D.=17.28) with significantly higher frequency than both Scanning (Mean=25.51 S.D.=18.12) (Observed t=-4.02, p<.01) and Detailed reading (Mean=31.02 S.D.=18.14) (Observed t=2.78, p<.05). There was no significant difference in frequency of Scanning and Detailed reading (Observed t=-1.16, n.s.). Figure 4.4 shows the mean frequency (in percentage of time) for Scanning, Skimming and Detailed reading.
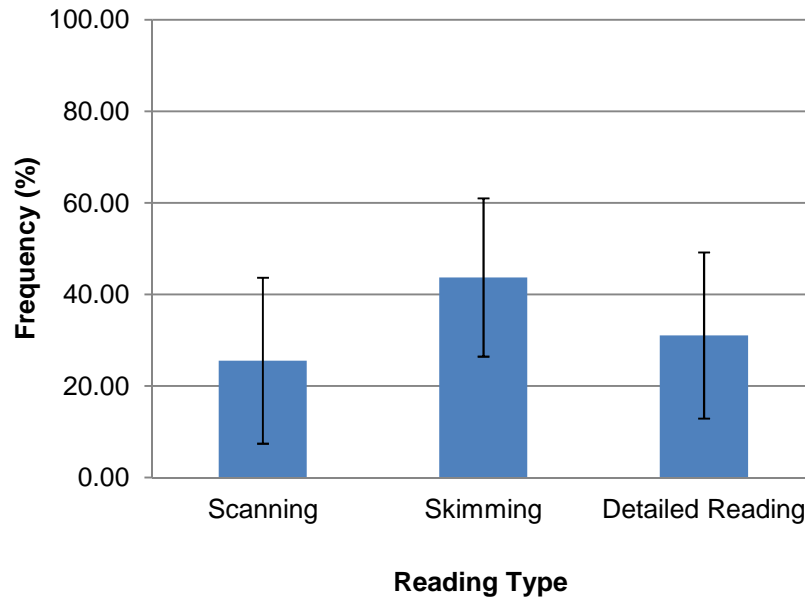
Figure 4.4: Mean frequency for Scanning, Skimming and Detailed reading

For the percentage of time which participants believe the spend using each reading type, an ANOVA found that the main effect for Reading Type was significant, ($F_{(2,88)}$ = 13.38, $p<.01$, $\eta_p^2 =.23$). A Scheffé post hoc found that participants estimated that they spent significantly more time Scanning (Mean=20.22 S.D.=17.93) than both Skimming (Mean=36.00 S.D.=15.87, Observed $t =-3.86$, $p<.01$) or Detailed reading (Mean=43.78 S.D.=19.89, Observed $t =-4.60$, $p<.01$). There was no significant difference between Skimming and Detailed reading (Observed $t =-1.17$, n.s.). Figure 4.5 shows the mean percentage of time which participants estimated they spent using Scanning, Skimming and Detailed reading.
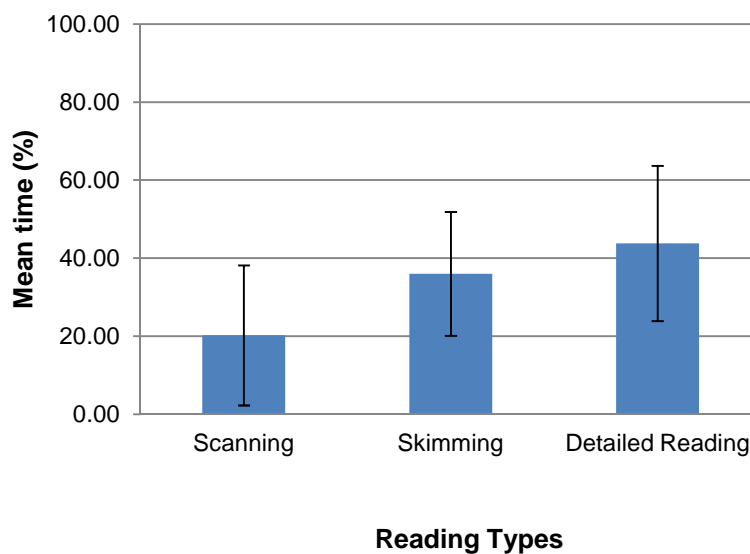
Figure 4.5: Mean time (percentage) spent using Scanning, Skimming and Detailed reading

4.6 Discussion

The results of this study provide evidence that the type of reading plays a part in both the accuracy of answers to questions and time taken to do the reading. While scanning accuracy relates to the ability to find a word, both skimming and detailed reading relate to the comprehension of the text. In these cases, while skimming is much faster than detailed reading this comes at a trade-off of lower accuracy. In addition, the results show that skimming provides a similar advantage as that of scanning in comparison to detailed reading, in that it produces lower fatigue both physically and visually.

However, there is an interesting issue with the results of this study in that line spacing did not seem to have any effect on the objective performance measures. This is odd, as Ling and van Schaik (2007) who did the research with the same line spacing condition as the current study found that visual search, a scanning task, did produce a substantil difference in performance for different line spacings.

In trying to understand the failure to replicate Ling and van Schaik's results, I realized that the task used in my study was not exactly the same as the visual search task used in Ling and van Schaik (2007). In particular, the my study asked participants to find a target word in the text of the webpage, whereas Ling and van Schaik (2007) asked participants to find a hypertext link word.

In order to see if this difference in the task played a role in the lack of a significant difference in the objective measures, further data was collected with the same participants (as many of those who were available to participate in a further session) with a task that more accurately replicated the Ling and van Shaik (2007) one, in order to identify if this was in fact the reason for the difference in results. The new task will referred to as *Searching for Link Word*.

4.7 Additional data collection

4.7.1 Experimental Design

The additional data collection had a two way mixed design. Line Spacing was the between participants independent variable. Reading Type was within participant independent variable. The independent variables had the following levels:

- Line Spacing - Single Line Spacing, 1.5 Line Spacing, or Double Line Spacing
- Reading Type - Seaching for link word (target present), or Searching for link word (target absent)

Each participant undertook six tasks, three tasks with each reading type. Order of reading type and the text were counterbalanced between participants to avoid fatigue and practice effects. The dependent variables were Time spent per webpage, Percentage of correct answers, Participants' ratings of visual and physical fatigue, and Participants' perceptions of their use of the reading types questionnaire which all mentioned in Section 4.4.1.

The data from the additional data collection were added to the data already collected. As there are two tasks for each of Reading Type in the main study and three tasks for each of additional data collection. Therefore the average time and percentage of correct answers in each Reading Type condition was calculated for further analysis.

4.7.2 Method

4.7.2.1 Participants

There were 15 participants, 11 males and 4 females, from the first round of data collected who re-participated. The mean of age was 27.1 years (SD=4.3, range 19 - 34

years). Two participants were undergraduate students, 8 were Ph.D. students and 3 were employed.

5 participants re-participated in the Single Line Spacing condition, 4 re-participated in the 1.5 Line Spacing condition, and 6 re-participated in the Double Line Spacing condition.

4.7.2.2 Equipment and materials

Equipment and materials in the additional data collection were the same as in the main experiment, with the following changes.

A new website with three pages for a practice task and six pages for the experimental tasks was created for this additional data collection. Each page had 275 words of text about the Olympic Games, separated into four paragraphs. For each text, seven nouns of approximately the same length to one another, distributed throughout the text, were marked up as hypertext links.   The mean number of characters for the link words was 8.88 (SD=1.56, range: 7–12 characters). In the "present" condition, the target links were presented in the upper-left, middle-right, or bottom-centre of the webpage an equal number of times. In the "absent" condition, no target link word appeared on the webpage. In three practice trials, the target link word was present twice and absent once. In six pages for experimental, the target link word was present four times and absent two times.

All texts and hypertext links on webpages were Arial 12 point with 1.5 line spacing and left only justified. The texts on the webpages were always black text (#000000) on white background (#FFFFFF) while hypertext links were presented in blue text (#0000FF) and underlined. As before, three versions of the website were created corresponding to the three levels of Line Spacing; Single, 1.5, Double Line Spacing.

4.7.3 Procedure

The procedure was similar to the first data collection, with the following differences.

Participants were given onscreen instructions explaining that they were going to perform a *Searching for link word* task, specifically that they were to try to find a hyperlink word on a webpage as quickly and as accurately as possible. If the link was present on the page, they had to press the 'P' key, and if it was absent they had to press 'A'.

After reading the instructions, participants pressed the 'S' key on the keyboard to start the trials. For each trial the following sequence of screens were presented: a blank white screen for two seconds, then a target hypertext link word in Arial 48 point font at the centre of a white background for one second, two seconds of blank white screen, and finally an text on a webpage. The participant's response automatically initiated the next trial. If a participant had not responded after seven seconds then the next trial automatically started.

After participants completed the practice trails, they were asked to do the six Searching for Link Word task with the same Line Spacing which they had experienced in the previous session. Once the participants completed all the tasks, they were asked to complete the visual fatigue questionnaire and participants' perceptions of their use of the reading types questionnaire.

4.7.4 Data preparation

The data which collected from the additional data collection were added to the data already collected. As before, I did the histograms from the data on Time spent per webpage in each of Reading type. The histogram was not a normal distribution then it was necessary to normalise before doing the analysis by using the method outlined in Section 3.4.1. In this process, 3 from 75 data (4%) were adjusted.

4.7.5 Results

There was a different number of task in each Reading Type: three tasks in each Searching for link word condition (target present and absent conditions) and two tasks each in Scanning, Skimming, and Detailed reading. Therefore the average time and average percentage of correct answers in each Reading Type condition was calculated for further analysis. All times were measured in seconds.

4.7.5.1 Time spent per webpage

An Analysis of Variance (ANOVA) found that the main effect for Line Spacing was not significant ($F(2,12)=0.39$, n.s.). There was a significant main effect for Reading Type ($F(4,48)=64.72$, $p<.001$, $\eta_p^2 =.84$).There was no significant interaction between the variables.

A Scheffé post hoc analysis was used to investigate the specific differences between Reading Type.

Table 4.2: Observed t - values between all pairs of Reading Type for Time spent per webpage

	Searching for link word (target present)	Searching for link word (target absent)	Scanning	Skimming	Detailed Reading
Searching for link word (target present)	-	3.88	-9.27**	-10.52**	-8.80**
Searching for link word (target absent)	-	-	-8.62**	-10.33**	-8.69**
Scanning	-	-	-	-5.75**	-6.75**
Skimming	-	-	-	-	-7.84*
Detailed Reading	-	-	-	-	-

* p < 0.05 ** p < 0.01

A Scheffé post hoc found that each reading type was different from each other, except Searching for link word (target present) (Mean=2.15 seconds, SD=1.04) which was not different from Searching for link word (target absent) (Mean= 3.09 SD=0.75, Observed t=-3.88, n.s.).

Searching for link word (target present) (Mean=2.15, SD=1.04) took a significantly shorter time than Scanning (Mean=19.43, SD = 7.45, Observed t = -9.27, p<.01), Skimming (Mean=41.96, SD=14.71, Observed t=-10.52, p<.01), and Detailed Reading (Mean=75.27, SD=32.50, Observed =-8.80, p<.01).

In addition, Searching for link word (target absent) (Mean=3.09, SD=0.75) took a significantly shorter time than Scanning (Mean=19.43, SD=7.45) (Observed t =-8.62, p<.01), Skimming (Mean=41.96, SD=14.71) (Observed t=-10.33, p<.01), and Detailed Reading (Mean=75.27, SD=32.50) (Observed t =-8.69, p<.01).

Finally, Scanning (Mean=19.43, SD=7.45) took a significantly shorter time than Skimming (Mean=41.96, SD=14.71) (Observed t=-5.75, p<.01), and Detailed Reading (Mean=75.27, SD=32.50) (Observed t =-6.57, p<.01). Skimming (Mean=41.96, SD=14.71) took a significantly shorter time than Detailed Reading (Mean=75.27, SD=32.50, Observed t =-4.84, p <.05).

Figure 4.6 shows the mean time per webpage (sec) for Searching for link Word (target present), Searching for link Word (target absent), Scanning, Skimming, and Detailed reading.
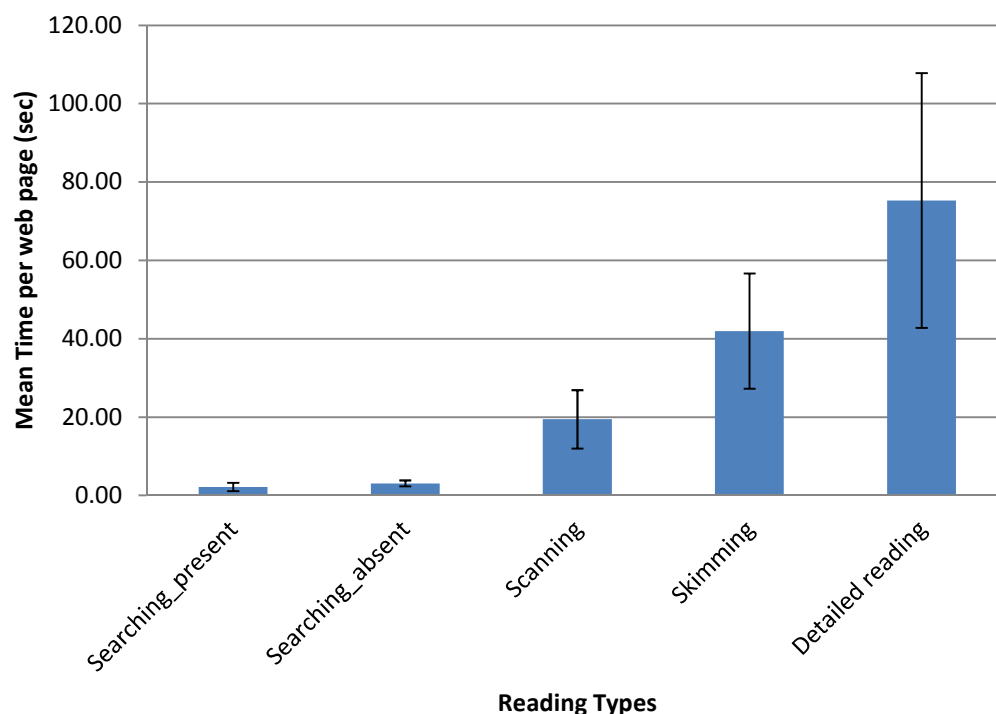
Figure 4.6: Mean Time per Webpage (sec) for Searching for link word (target present and absent), Scanning, Skimming, and Detailed reading

4.7.5.2 Percentage of correct answers

An ANOVA found that the main effect for Line Spacing on percentage of correct answers was not significant ($F_{(2,12)}=0.15$, n.s.). There was a significant main effect for Reading Type, ($F_{(4,48)}=20.87$, $p<.001$, $\eta_p^2=.64$). There was no significant interaction between the two variables.

A Scheffé post hoc found that Skimming (Mean 42.50 S.D.=19.93) produced a significantly lower percentage of correct answers than Searching for link word (target present) (Mean=86.67 S.D=30.34, Observed t=5.61, p<.01), Searching for link word (target absent) (Mean=84.44 S.D.=30.52, Observed t= 5.42, p<.01), Scanning (Mean=100 S.D.=0, Observed t=11.17, p<.01), and Detailed Reading (Mean=71.67 S.D.=22.39, Observed t=-4.85, p<.05). Finally, Detailed reading (Mean=71.67 S.D.=22.39) produced a significantly lower percentage of correct answers than

Scanning (Mean=100 S.D.=0, Observed t=4.94, p<.05). There were no other significant differences.
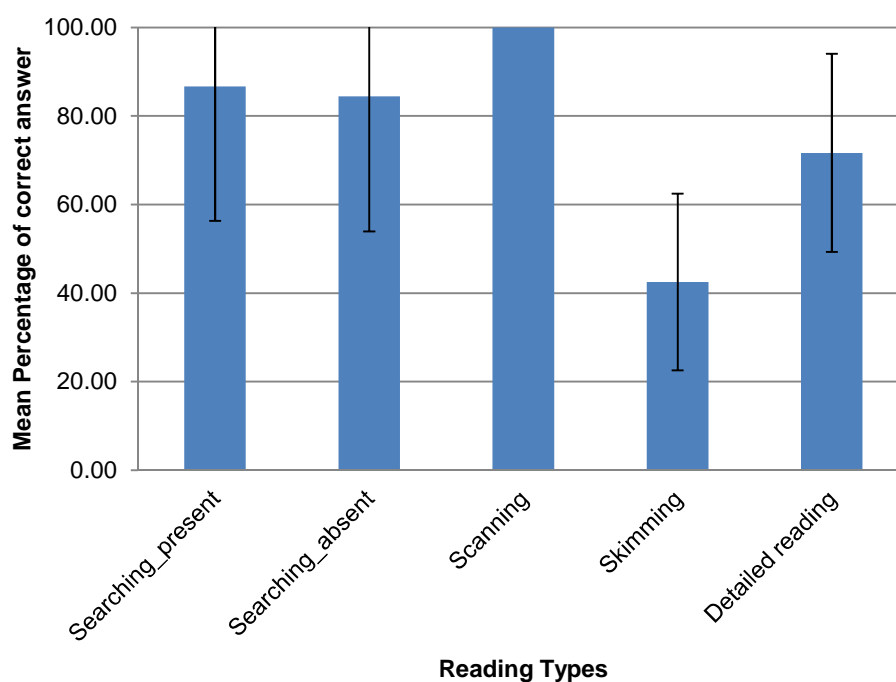
Figure 4.7: Mean Percentage of correct answers for Searching for link word (target present and absent) Scanning, Skimming, and Detailed reading

Figure 4.7 shows the mean percentage of correct answers for Searching for link Word (target present and absent), Scanning, Skimming, and Detailed reading.

4.7.5.3 Visual and physical fatigue

The Visual and Physical Fatigue were measured the levels of visual and physical fatigue in five aspects (5 questions) and the last question for the overall fatigue on a 7 point Likert items. Correlation analyses between the participants' rating on these five aspects (question 1-5) and overall fatigue (question 6) were assessed to investigate whether the participants were measuring the different things.

This strong pattern of correlations, see Appendix 19, suggests that the overall fatigue in question 6 represented the other 5 questions of visual and physical fatigue. The overall fatigue for each participant for each of the reading typre was used for the further analysis.

An ANOVA found that the main effect for Line Spacing on the Overall Fatigue ratings was not significant ($F_{(2,12)}=0.02$, n.s.). There was a significant main effect for Reading Type ($F_{(3,36)}= 8.33$, $p<.001$, $\eta_p^2 =.41$). There was no significant interaction between the variables.

A Scheffé post hoc found that Detailed reading (Mean=3.73 S.D.=1.28) was rated significantly higher in overall fatigue than both Searching for link word (Mean=2.40 S.D.=1.30, Observed t=-4.39, p<.05), and Skimming (Mean=2.67 S.D.=1.40, Observed t=-4.30, p<.05). There were no other significant differences.

Figure 4.8 shows the mean Overall Fatigue for Searching for link Word (both target present and absent), Scanning, Skimming, and Detail reading.
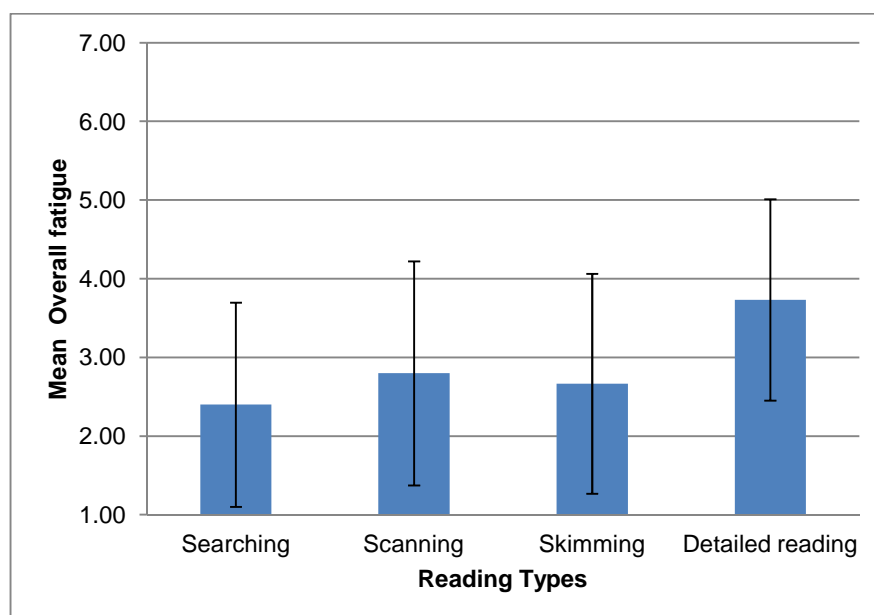
Figure 4.8: Mean Overall Fatigue ratings for Searching for link word (both target present and absent) Scanning, Skimming, and Detailed reading

One sample t-tests against the neutral mid-point rating of 4 showed that the mean ratings for Searching for link word, Scanning and Skimming were significantly lower than neutral (Searching for link word: $t_{(14)}=-4.77$, $p<.001$; Scanning: $t_{(14)}=-3.26$, $p<.01$; Skimming: $t_{(14)}=-3.70$, $p<.01$). While the mean rating for Detailed reading was not significant different from neutral (Detailed reading: $t_{(14)}=-0.81$, n.s.).

4.8. Discussion and Conclusions

In spite of collecting and analysing add additional data using the Ling and van Shaik (2007) methodology, line spacing was not shown to have a significant effect on

performance on the Searching for Link Word task. The results differ from Ling and van Shaik (2007) who used the Searching for Link word task and found a significant effect of line spacing on accuracy and reaction times. However, for current study, line spacing had no effect on participants' performance on any reading task types used: Searching for Link Word, Scanning, Skimming, and Detailed Reading.

O'Hara (1996) noted that web users use different types of reading for different purposes. Information on the web may be skim read, scanned for a piece of information, or read for comprehension. Using such a range of reading types in an experiment would be good as it reflects the real range of reading activities of web users. However, it would create a very complex and time consuming study for participants and mean that fewer variables could be investigated at any one time. These issues are of particular concern when doing research with older participants. It would preferable to use a single type of reading task to overcome these issues.

When considering appropriate reading tasks to be used in the subsequent experiments about reading text on webpages in this programme of research, it is important to select a reading task which is ecologically valid (O'Hara and Sellen, 1997; Pearson and van Schaik, 2003). In order to select a single reading type for the subsequent studies, there are five dimensions to be taken into consideration in making a decision: how time consuming a task is; how much overall fatigue it puts on participants, particularly relevant for older participants; the ability of the task to generate different reading speeds and thus allow discrimination in reading performance; participants' feelings in doing the task, as it is unethical to subject participants to unnecessary upset; and finally the ecological validity of the task in relation to web users' actual behaviour.

In terms of the time consuming nature of the task, detailed reading required a signficantly longer time on each webpage over the other reading types. Searching for link word and scanning took much shorter times while skimming took a moderate time. For this dimension therefore, the tasks with the shorter times, searching for link word, scanning, and skimming, are recommended.

In terms of overall fatigue, detailed reading produced the highest level of overall fatigue, significantly higher than skim reading searching for a link word and scanning, which were not significantly different from each other. Therefore, for this dimension, searching for link word, scanning, or skim reading, are recommended.

In term of producing a high percentage of correct answers, participants made some erroneous answers on every reading type, except scanning. In the scanning task, all

participants produced the same completely successful rate (100%) regardless of line spacing condition. Thus, if using the scanning task in further studies, the independent variables may produce a similar ceiling effect, and there will be no variation in the percentage of correct answers. For the other reading types participants had different percentages of correct answers, so the independent variables have the possibility to produce significant effects. For this dimension, searching for link word, skimming, and detailed reading are recommended.

In term of participants' feeling, many participants gave feedback after completing the four different reading tasks that they were nervous and stressed while doing the searching for link word and scanning tasks. In the searching for link word task, participants had only seven seconds for complete each trial. If they did not respond within seven seconds then the next trial automatically started, which made some participants felt nervous and stressed. For the scanning task, there was no time limit for completing each trial, but the nature of the task is quite similar to searching for a link word. Some participants were also stressed during the scanning task as they felt that they had to complete each trial very quickly. Sometimes when the participants completed the trails slower than they had expected, they felt nervous. However, participants did not feel particularly nervous when undertaking the skim reading and detailed reading tasks. Thus, for this dimension, skimming and detailed reading are recommended.

In terms of being ecologically valid, as mentioned above, all the reading types are ecologically valid but which reading task reflects web users' most typical behaviour? Skim reading is the reading type that web users reported that they used more frequently when they read text on webpages. In addition, participants reported using skim reading and detailed reading for the same percentage of time. Thus for this dimension, skim reading is recommended due to the results on both estimated frequency of use and estimated amount of time used.

In terms of the limitations of the study, the participants in this current study were all younger adults. Older participants were not recruited as the study was complex and time-consuming, including a number of different reading tasks. Participants needed to switch between the different reading types which it was not particular hard for younger adults but might have been stressful for older adults. So, if the older adults participated in the study, the results might longer time or higher overall fatigue than the current results.

Conclusion: an appropriate task for research about reading webpages

The discussion on the five dimensions above shows that skim reading is not time consuming and participants did not become nervous when doing it because of any time limit in completing the task. In addition, skim reading produced less overall fatigue but showed a variations in the percentage of correct answers. Finally, skim reading reflected how web users estimate their use of read tasks on the web both in terms of frequency and percentage of time used. Thus, I chose the skim reading task as the reading task for use in the subsequent experiments in this programme of research.

Chapter 5

Effects of font type and font size on skim reading webpages by younger and older people in the UK and Thailand

---

5.1 Introduction

This chapter presents the results of the third experiment in this programme of research which investigated the effect of font type and font size on skim reading webpages by younger and older adults both in Thailand and the UK.

Font type and font size were chosen to investigate as independent variables in the third experiment because both of these aspects of text presentation are mentioned in multiple sets of guidelines on web accessibility for older adults, as discussed in the literature review in Chapter 2. However, the recommendations in these guidelines usually not supported by evidence and are unclear as to why particular presentations would be better for older adults.

Font type is one of aspects of text presentation which many web design guidelines mention but the recommendations they make about this feature are varied. In seven sets of guidelines which mention font type, recommendations about font type are: Zhao (2001) recommends choosing font type based on "legibility" without explaining what that means; Agelight (2001) similarly suggests choosing font type based on "familiarity and legibility"; the SPRY Foundation (Holt and Komlos-Weimer, 1999), Holt (2000), SilverWeb (Kurniawan and Zaphiris, 2005; Zaphiris, Kurniawan and Ghiawadwala, 2007), the National Institute on Aging/National Library of Medicine (NIA/NLM, 2002), and ARRP (2004) all recommend using a sans serif font type; SilverWeb (2005, 2007) warns against the use of fancy font types; and NIA/NLM (2002) warns against the use

of serif, novelty, and display font types; ARRP (2004) refers to the fact that some research suggested serif fonts for speed but that readers prefer sans serif fonts.

The recommendations about font size provided by the guidelines are also varied. WebCredible (Fidgeon, 2006) suggests that less than 12 point is too small to read but do not recommend a specific font size; Holt (2000) also does not provide specific font size. Zhao (2001) recommends at least 12 point; Agelight (2001), the SPRY Foundation (Holt and Komlos-Weimer, 1999) and SilverWeb (Kurniawan and Zaphiris, 2005; Zaphiris, Kurniawan and Ghiawadwala, 2007) recommend 12 to 14 point; while NIA/NLM (2002) recommends 12 or 14 point.

Currently, there is no specific recommendations, nor evidence regarding user performance or preference, on font type and font size for Thai language web sites in any web design guidelines.

This experiment tested a range of combinations of font type and font size with younger and older participants from both the UK and Thailand.

The following research questions were addressed by this study:

> 1. Does font type have an effect on reading performance and preferences of younger and older people when reading web pages?

> 2. Does font size have an effect on the reading performance and preferences of younger and older people when reading web pages?

> 3. Does age group have an effect on reading performance and preferences when reading web pages?

> 4. Does nationality have an effect on reading performance and preferences when reading web pages?

> 5. Do attitudes toward the web have an effect on reading performance when reading web pages?

5.2 Font Types in Latin and Thai alphabets

To investigate font type in English and Thai texts, decisions needed to be made not only about the font types to use for the English texts, but also for the Thai texts. This section explains fonts in the Thai alphabet and their equivalent to fonts in the Latin alphabet. The Thai Conservative font type includes an extra circle at the beginning of

most of consonants and vowels (Figure 5.1 and Figure 5.2), while the Modern font type does not (Figure 5.3 and 5.4). Thus, the Thai Conservative font type corresponds most closely to a serif font in the Latin alphabet and the Thai Modern font type correspond most closely to a sans serif font.

กีฬาโอลิมปิกเป็นงานแข่งขันกีฬาระดับนานาชาติครั้ง
สำคัญที่ครอบคลุมทั้งกีฬาฤดูร้อนและกีฬาฤดูหนาว
มีนักกีฬาหลายพันคนเข้าร่วมในการแข่งขันกีฬาหลาก
หลายประเภท   โอลิมปิกเกมส์ได้รับการยอมรับว่าเป็น
งานแข่งขันกีฬาครั้งสำคัญที่สุดของโลก

Figure 5.1: An example of text written in the Thai Conservative font.

serif สรชัย

Figure 5.2: An example of a Latin alphabet serif font and a Thai alphabet conservative font. The red circled areas show extra annotation on the characters.

กีฬาโอลิมปิกเป็นงานแข่งขันกีฬาระดับนานาชาติครั้ง
สำคัญที่ครอบคลุมทั้งกีฬาฤดูร้อนและกีฬาฤดูหนาว
มีนักกีฬาหลายพันคนเข้าร่วมในการแข่งขันกีฬาหลาก
หลายประเภท   โอลิมปิกเกมส์ได้รับการยอมรับว่าเป็น
งานแข่งขันกีฬาครั้งสำคัญที่สุดของโลก

Figure 5.3: An example of a text written in Thai Modern font.

sans serif สรชัย

Figure 5.4: An example of a Latin alphabet sans serif font and a Thai alphabet modern font. The red circled areas show areas absent of extra annotation.

5.3 Method

5.3.1 Design

A four way mixed design was used in this experiment. Age Group and Nationality were the between participant independent variables, and font size and font type were the within participant independent variables.

The independent variables had the following levels:

- Age Group - participants were either Older Adults (55 years and over for participants in Thailand and 65 years and over for participants in the UK, see Section 2.3 for calculation of appropriate minimum age for older adults in Thailand and the UK) or Younger Adults (18 to 39 years in both Thailand and the UK)

- Participants nationality, language and writing system (Nationality for short) - participants were either British people in the UK who were native speakers of English or Thai people in Thailand who were native speakers of Thai

- Font Size – either 12, 14, or 16 point

- Font Type – either serif: Times New Roman and Thai Conservative font type or sans serif: Arial and Thai modern font type

Each participant undertook six tasks on web pages, one task with each of the six combinations of Font Type and Font Size. Tasks were to skim read the text and answer four multiple choice questions. The order of presentation of the six combinations was counterbalanced between participants to compensate for practice and fatigue effects.

Two dependent variables related to performance were measured:

- Time spent per web page – this was a measure of the reading speed of participants on a web page.

- Number of correct answers – this was measure of the efficiency. The number of correct answers were calculated and converted to the percentage of correct answer before analyses were undertaken

Five dependent variables related to participants' preferences were measured:

- Participants' attitude towards the web were measured using the Attitudes toward the Web Scale (Burn, 2003) which has three factors: Confidence, Performance, and Fashion (see Chapter 2, section 2.4, for more detailed information)

- Participants' rating of visual and physical fatigue (adapted from Tyrrell and Leibowitz, 1990)

- Participants' ratings of each Font Type condition on three dimensions: Ease of Reading; Pleasantness of Reading; and Speed of Reading

- Participants' ratings of each Font Size condition on three dimensions: Ease of Reading; Pleasantness of Reading; and Speed of reading

- Participants' ratings of their overall preference of each of combination of Font Type and Font Size

Participants' ratings on the visual and physical fatigue was measured at the end of each task while the other ratings were measured when the participants had completed all the tasks. All the preference dependent variables were all measured on five point Likert items, except the visual and physical fatigue questionnaire which was measured on a seven point Likert items. The seven point Likert scale of visual and physical fatigue was adopted from Dillion, Kleinman, Ok Choi, and Bias (2006).

5.3.2 Participants

72 people participated in this experiment. 30 participants in the UK and 42 in Thailand.

The UK participants comprised 18 Younger and 12 Older Adults. The UK Younger Adults comprised 12 males and 6 females, aged between 18 and 33 years (Mean = 20.56 years, SD = 3.88). There were 16 undergraduate students, one Masters degree

student, and one Ph.D. student. All the Younger Adults had no experience with the previous experiments in this programme of research. The UK Older Adults comprised 6 males and 6 females, aged between 65 and 87 years (Mean = 73.75 years, SD = 7.15). All the UK Older participants were retired.

The Thai participants comprised 21 Younger and 21 Older Adults. The Thai Younger Adults comprised 8 males and 13 females, aged between 21 and 39 years (Mean = 27.71 years, SD = 4.92). Four were undergraduate students, four were Masters students, five were Ph.D. students and eight were employed. The Thai Older Adults comprised 11 males and 10 females, aged between 60 and 76 years (Mean = 61.67 years, SD = 3.32). 16 were employed and the other five were retired.

The Younger Adult participants were offered a gift voucher valued at £10. The Older Adult participants were offered a gift voucher valued at £15, as the sessions for the Older Adults took considerably longer than those for Younger Adults.

The UK Younger Adult participants were recruited by sending an email to students in Department of Computer Science, the University of York, and a message was also posted on the Uiversity of York Graduate Student Association web site.

The UK Older Adult participants were recruited from panel list of older people who had participated with previous studies for the Human-Computer Interaction Research Group in the Department of Computer Science and the York Older People's Assembly.

The Thai Younger participants were recruited from students, staff, and alumni of Suranaree University of Technology in Thailand. The Older participants were recruited from lecturers and staff of Suranaree University of Technology inThailand.

In both the UK and Thailand, a snowball recruiting strategy was also used with the Older Adult participants, once someone had taken part in the study, they were asked to ask their friends if they would like to participate in the study.

5.3.3 Equipment

The study was conducted on a personal computer (Acer Aspire 4741, Intel (R) Core (TM) i5) running Windows 7 Home Premium and Internet Explorer 9 with a standard keyboard, 2-button mouse with a scroll-wheel, video camera, and 21.5 inch LED

Monitor. The screen resolution was 1920 x 1080 pixels. Morae[1] software was used to record and analyse the sessions.

5.3.4 Materials

5.3.4.1 Pre-study questionnaires

There were two pre-study questionnaires. The first pre-study questionnaire was adapted from the Attitudes toward the Web Scale (Burn, 2003), a set of 18 statements about respondents' attitudes towards the web, which are rated on five point Likert items (1 = strongly disagree, 5 = strongly agree). This scale has three Factors: Confidence, Performance, and Fashion, see Chapter 2, Section 2.4 for more details. The full set of questions and scoring of the scales can be found in Appendix 10. The data collected from this questionnaire provided participants' attitude towards the web, which was one of the preference dependent variables.

The second pre-study questionnaire consisted of a set of questions about the participants' use of the web and demographic information. Questions included information about age, gender, occupation, experience with the web, and use of the web. The questions also asked participants to rate their level of computer experience (on a 7 point Likert item: 1 = none at all, 7= extensive) and their expertise in using the web (on a 7 point Likert item: 1 = none at all, 7=expert). The full set of questions can be found in Appendices 3 and 4, for the English and Thai versions respectively.

5.3.4.2 Post-study questionnaire

A post-study questionnaire measured participants' ratings of Ease, Pleasantness, and Speed of Reading for each condition of Font Type and Font Size, and their Overall rating of Preference of each combination of Font Type and Font Size on a 5 point Likert items (1 = least preferred / strongly disagree, 5 = most preferred / strongly agree). The questionnaire was adapted from the first and the second experiments. The questionnaire was similar to the questionnaire in Appendix 5 and 6, for the English and Thai versions respectively.

---

[1] http://www.techsmith.com/morae.html

The post-study questionnaire provided three preference dependent variables: participants' ratings of each Font Type on three dimensions, participants' ratings of each Font Size on three dimensions, and participants' ratings of their overall preference of each of the combination of Font Type and Font Size.

5.3.4.3 Visual and physical fatigue questionnaire

The visual and physical fatigue questionnaire was adapted from Tyrrell and Leibowitz (1990) while the 7 point Likert items (1 = strongly disagree; 7 = strongly agree) adopted from Dillion, Kleinman, Ok Choi, and Bias (2006). This questionnaire was the same as used in the second experiment, see Appendix 13 and 14, for the English and Thai versions respectively.

The visual and physical fatigue questionnaire provided the data for one of the dependent variable in this study.

5.3.4.4 Experimental web site

Two multi-page websites were created for this study, one in English and one in Thai. The English website was adapted from the materials discussed in the experiment detailed in Chapter 3. The Thai website was adapted from the experiment detailed in Chapter 3 and was checked for quality of translation as per the protocol given in Appendix 19. Each website had a page for practice task and six pages for the experimental tasks.

On the English website, each page had 275 words of text about the Olympic Games, separated into four paragraphs. The texts were approximately the same ease of reading. On the Thai website each page had 354 words of text, separated into four paragraphs. This number of words was approximatly the same length of text as 275 words in English. However, when the text, both English and Thai version, were separated into four paragraphs, there were some different lengths in paragraphs, as shown in Figure 5.5 and Figure 5.6. The content on web page in the study were similar to those found in Appendices 11 and 12, for the English and Thai versions respectively.

In line with the results of the first experiment, the content was organised to left - right justified with 1.5 line spacing.

Figure 5.5: Example page from the experimental web site (English version)
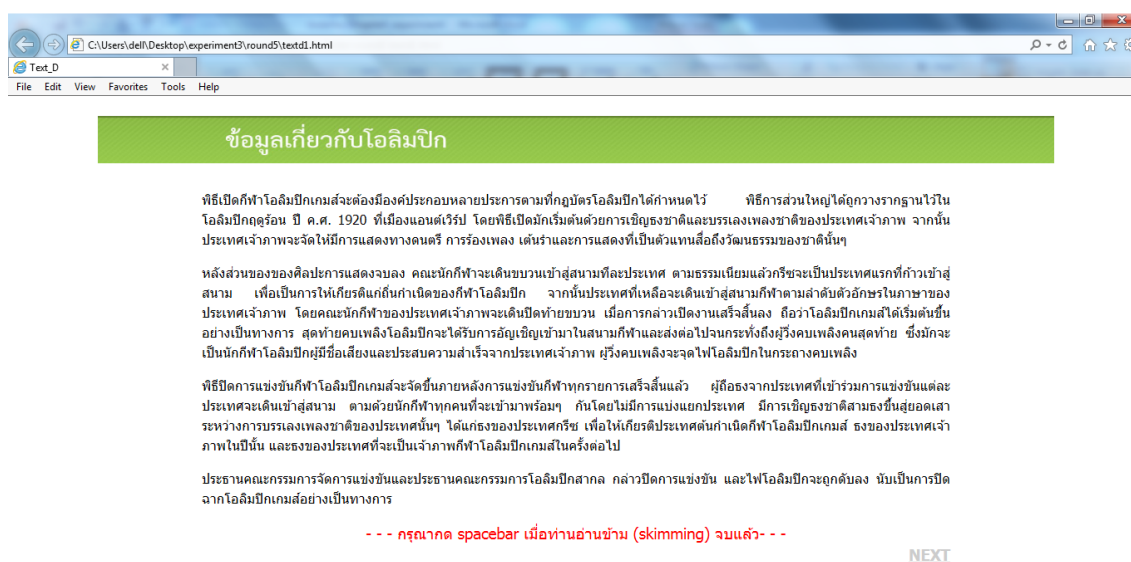
Figure 5.6: Example page from the experimental web site (Thai version)

5.3.4.5 Experimental tasks and multiple choice questions

The tasks asked participants to skim read the text on a web page as quickly and accurately as possible in order to understand the main ideas of the text. [The participants were informed to press the space bar to indicate that they had finished skim reading a text. Then, participants were asked to complete four multiple choice questions.

After each task there was a set of four multiple choice questions. All multiple choice questions could be answered from content found in the text and had the approximatly the same level of difficulty. The last choice of each multiple choice question was "I'm not sure" for avoiding participants answering the questions correctly by guessing. All multiple choice questions were printed on paper to avoid visual fatigue during completing the questions. Because the task is reading, which was done from the screen. The full set of multiple choices questions were found in Appendices 11 and 12, for the English and Thai versions respectively.

5.4 Procedure

The study was conducted at a number of locations, all quiet rooms at the institutions where the participants studied or where they came to take part in the study.

Before starting the study, participants were briefed about its nature, and their rights. Any questions that participants had about the study were answered. When they were happy about participation in the study, they were asked to sign the consent form section A (see Appendices 1 and 2, for the English and Thai versions respectively). After signing the consent form, participants were asked to complete the pre-study questionnaire (see Appendices 3 and 4, for the English and Thai versions respectively).

Participants were asked to sit approximatly 57 cm. from the monitor that is the average distance which most people feel comfortable (College of Optometrists, 2013). They were then asked to use the web browser to explore the web to make themselves familiar with the computer, monitor, mouse, and the Internet Explorer web browser. When participants were comfortable and ready to start, the participants were given a practice task to familiarise themselves with the task required in the study.

For each experimental task, participants skim read the text on a web page with the appropriate combination of font type and size as quickly and accurately as possible in order to understand the main ideas of the text. The participants were asked to press the space bar to indicate that they had finished skim reading on the text. Then, participants were asked to complete the four multiple choice questions about the text and the visual and physical fatigue questionnaire. The multiple choice questions and the questionnaire were presented on paper to avoid further visual fatigue. Once the participants completed these items, they were asked to continue to the next task by clicking the "Next" button when they were ready. On clicking "Next" a new text with a different combination of Font Type and Font Size was shown to the participants. This

process was repeated for each of the six tasks. All task sections of the experiment were recorded using Morae for later analysis.

The order of presentation of the web pages with the six combinations of Font Type and Font Size was counterbalanced between participants.

After completing all the tasks, participants were asked to rate the Ease, Pleasantness, and Speed of Reading for each condition of Font Type and Font Size, and their overall rating of preference of each of combination of Font Type and Font Size using the post-study questionnaire (similar to Appendix 5 and 6). As a reminder, examples of Font Type, Font Size, and six combinations were provided to participants.

Participants were then debriefed and the purpose of the study was fully explained to them and any questions they had were answered. Participants were then asked to sign the Section B of consent form to show they were happy with their experience.

Each session took approximately 20 minutes to complete for Younger Adult participants, and approximately 40 minutes to complete for Older Adult participants.

5.5 Data preparation

Firstly, I did the histograms from the Data on Time spent per webpage in each combinations of Font Type and Font Size. The histogram was not a normal distribution, it was necessary to normalise before doing the analysis by using the method outlined in Section 3.4.1. In this process, 16 data points out of a total of 432 data points (3.70%) were adjusted.

5.6 Results

Analysis of Variance (ANOVA) was used to investigate the effects of the independent variables of Font Type and Font Size and the appropriate post-hoc analyses were conducted when there were any significant effects from the overall ANOVA analysis. In addition, for participants' preference ratings, t-tests were used to investigate whether preference ratings were significantly above or below the mid-point of the rating scale.

5.6.1 Time spent per web page

A four way Analysis of Variance (ANOVA) found that Font Type, Age Group and Nationality all had significant effects on the time spent per web page (Font Type: F

(1,68) = 6.26, p < 0.05, η_p^2 = 0.08; Age Group: F(1,68)=5.25, p<.05, η_p^2 =.07; Nationality: F (1,68)=22.76, p<.001, η_p^2 =.25). However, there was no significant effect of Font Size (F(2,136)=.47, n.s.). In addition, There were significant interactions between Font Type and Nationality (F(1,68)=7.19, p<.01, η_p^2 =.10), and between Age Group and Nationality (F(1,68)=9.52, p<.01, η_p^2 =.12). There were no other significant interactions between any of the independent variables.

The Serif fonts produced a significantly shorter skim reading times per web page than the sans serif fonts (Mean serif=69.77 sec. SD=25.97; Mean sans serif=73.35 sec. SD=30.24).

Older Adults spent significantly longer time skim reading per web page than Younger Adults (Mean Younger Adults=66.34 SD=29.86; Mean Older Adults=77.73 SD=24.81).

The UK participants spent significantly shorter time per web page than Thai participants (Mean UK participants=55.78 SD=18.93; Mean Thai participants=82.82 SD=28.37).

A Scheffé post hoc analysis was used to test the specific differences in the interaction between Font Type and Nationality. Table 5.1 shows the pattern of observed t-values for the two way interaction of Font Type and Nationality. For the UK participants, there was no significant difference on reading time between serif and sans serif (Mean serif=55.77 sec. SD=19.44; Mean sans serif=55.79 sec. SD=18.52) (Observed t = -.16, n.s. , critical t at 95% confidence level: 3.46, at 99%: 4.59). For Thai participants, serif was marginally read faster than sans serif font type. (Mean serif=79.76 sec. SD=25.47; Mean sans serif=85.89 sec. SD=30.79) (Observed t =-3.37, n.s.).

However, there were significant differences between the UK participants and the Thai participants reading time for both font types. The UK participants read serif text (Mean serif for UK participants=55.77 sec. SD=19.44) significantly faster than both Thai participants reading serif text (Mean serif for Thai participants=79.76 sec. SD=25.47) (Observed t=-7.51, p<.01), and Thai participants reading san serif text (Mean sans serif for Thai participants=85.89 sec. SD=30.79, Observed t=-8.16, p <.01). The UK participants also read sans serif text (Mean sans serif for UK participants=55.79 sec. SD=18.52) significantly faster than both Thai participants reading serif text (Mean serif for Thai participants=79.76 sec. SD=25.47, Observed t=-8.24, p<.01), and Thai participants reading sans serif text (Mean sans serif for Thai participants=85.89 sec. SD=30.79, Observed t=-8.24, p<.01). This interaction is shown in Figure 5.7.

Table 5.1: Observed t - values between all pairs of Font Type and Nationality combination for Time spent per web page

		UK		Thai	
		Serif	**Sans serif**	**Serif**	**Sans serif**
UK	Serif	-	-0.16	-7.51**	-8.16**
	Sans serif		-	-8.24**	-8.24**
Thai	Serif	-		-	-3.37
	Sans serif	-			-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.46, at 99%: 4.59
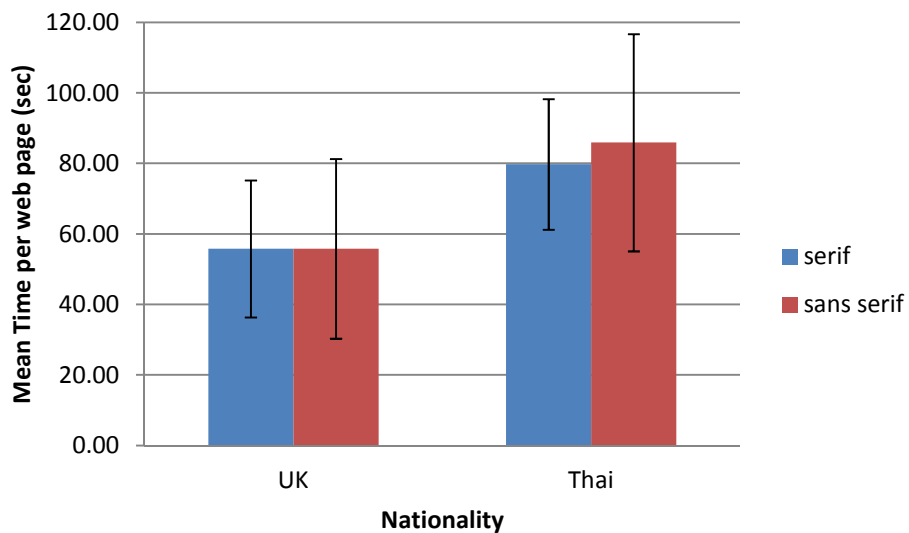
Figure 5.7: Mean Reading Time per web page (seconds) for serif and sans serif for the UK and Thai participants

A Scheffé post hoc analysis was used to test the specific differences in the interaction between For the interaction between Age Group and Nationality. Table 5.2 shows the observed t-values between each pair of combinations of Age Group and Nationality.

UK Younger Adults (Mean UK Younger Adults=44.80 sec. SD=8.42) read significantly faster than UK Older Adults, Thai Younger Adults, and Thai Older Adults (Mean UK Older Adults=72.25 sec. SD=18.40; Mean Thai Younger Adults=84.79 sec. SD=29.28; Mean Thai Older Adults=78.74 sec. SD=27.40; Observed t UK Younger Adults vs UK Older Adults=-11.38, p<.01; Observed t UK Younger Adults vs Thai Younger Adults=-

14.43, p<.01; Observed t UK Younger Adults vs Thai Older Adults=-12.56, p<.01) (critical t at 95% confidence level: 3.46, at 99%: 4.59). There were no significant difference between other combinations of variables. This interaction is shown in Figure 5.8.

Table 5.2: Observed t-values for all pairs of Age Group and Nationality combination for Time spent per web page

		UK		Thai	
		Younger	**Older**	**Younger**	**Older**
UK	Younger	-	-11.38**	-14.43**	-12.56**
	Older		-	-2.60	-1.48
Thai	Younger	-		-	1.21
	Older	-			-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.46, at 99%: 4.59
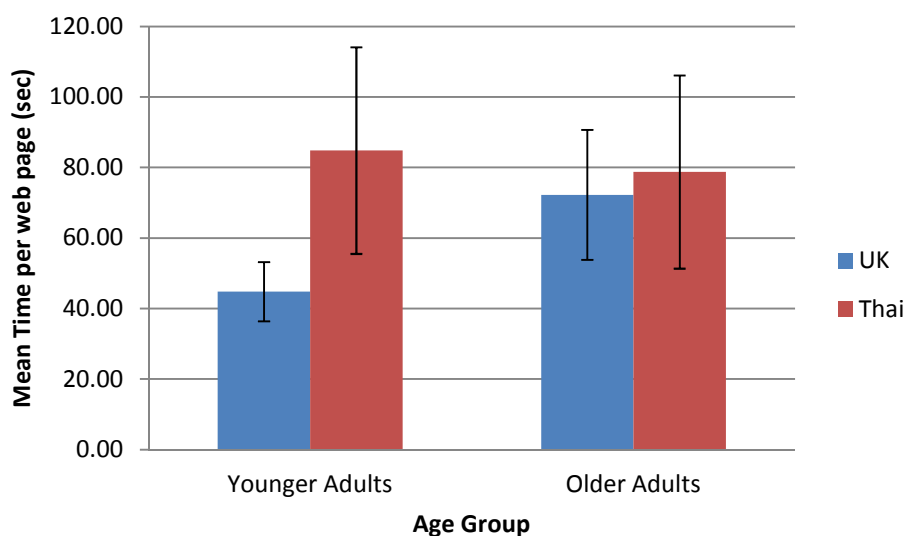
Figure 5.8: Mean Reading Time per web page (seconds) for Younger and Older Adults for the UK and Thai participants

5.6.2 Percentage of Correct Answers

For Percentage of correct answers, an four way Analysis of Variance (ANOVA) found that Font Type, Font Size and Age Group had no significant effects (Font Type: $F_{(1,68)}=.46$, n.s.; Font Size: $F_{(2,136)}=2.64$, n.s.; Age Group: $F_{(1,68)}=3.70$, n.s.).

However, there was a significant effect of Nationality (F(1,68)= 7.41, p<.01, η_p^2 =0.10). There were no significant interactions between any of the independent variables.

The UK participants were significantly more accurate at answering the questions than Thai Participants (Mean UK participants=53.19 SD=26.09; Mean Thai participants=44.84 SD=26.03).

To investigate whether there was a speed-accuracy trade-off in the way the participants undertook the tasks, the Time Spent per Web Page and the Percentage of Correct Answers were correlated. However, correlation were not significant for either Younger and Older Adults (Younger Adults: r(39)= -.13, n.s., Older Adults: r(33)=.19, n.s.).

5.6.3 Visual and Physical Fatigue

Visual and Physical Fatigue was measured with five questions and an additional question for overall fatigue on 7 point Likert scales. Correlation analyses between the participants' rating on these five questions and overall fatigue were assessed to investigate whether the participants were measuring different things.

This strong pattern of correlations (see Appendix 15), suggests that the overall fatigue in question 6 represented the other 5 questions of visual and physical fatigue. Therefore, the overall fatigue for each participant for each of the two Font Type levels and the three Font size levels was used for the further analysis.

A four way Analysis of Variance (ANOVA) found that Font Type, Age Group, and Nationality had no significant effect on the overall fatigue ratings (Font Type: F(1,68)=2.31, n.s.; Age Group: F(1,68)=2.53, n.s.; Nationality: F(1,68)=.10, n.s.). However, there was a significant effect of Font Size (F(2,136)=6.61, p<.01, η_p^2 =.09). There were no significant interactions between any of the independent variables.

A Scheffé post hoc analysis was used to investigate the specific differences between Font Size which is illustrated in Figure 5.9. Overall fatigue for 12 point text was significantly higher than overall fatigue for 16 point (Mean overall fatigue for 12 point=2.47 SD=1.56; Mean overall fatigue for 16 point=2.13 SD=1.33), Observed t-=3.65, p<.01, critical t at 95% confidence level: 2.47, at 99%: 3.09). There was no difference between Mean overall fatigue for 12 point and overall fatigue for 14 point (Mean overall fatigue for 12 point=2.47 SD=1.56; Mean overall fatigue for 14 point=2.28 SD=1.40), and between overall fatigue for 14 point and 16 point (Mean

overall fatigue for 14 point=2.28   SD=1.40; Mean overall fatigue for 16 point=2.13 SD=1.33, Observed t-values=2.03, n.s.).
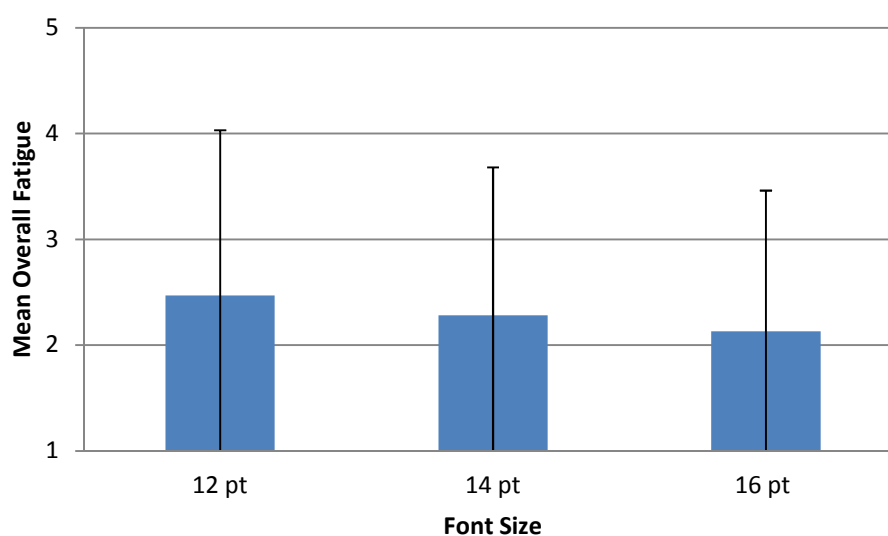
Figure 5.9: Mean overall fatigue for different Font Sizes

One sample t-tests of the overall fagitue ratings against the neutral mid-point rating of 4 showed that the mean ratings for the 12 point, 14 point, and 16 point were significantly lower than neutral (12 point: $t(143)=-11.79$, $p<.001$; 14 point: $t(143)=-14.68$, $p<.001$; 16 point: $t(143)=-16.96$, $p<.001$).

5.6.4 Preference measures

The preference measures were participants' ratings of Ease, Pleasantness, and Speed of Reading for each condition of Font Type, Font Size and their overall ratings of preference of each combination of Font Type and Font Size.

Participants were asked to rate three dimensions on 5 point Likert scales: Ease of reading, Pleasantness of Reading, and Speed of Reading. Correlation analyses between the participants' rating on these three dimensions were assessed to investigate whether the participants were measuring different constructs with these three ratings.

The strong pattern of correlations (see Appendix 16), suggests that participants had only one underlying experience dimension on which to rate the reading tasks. Therefore a combined User Reading Experience score (URE) was calculated for each participant for each of the two Font Type levels and the three Font Size levels and; this

URE scores was the mean of the three ratings for each condition. An added benefit of using this combined score is that scores made up from a number of individual measures are more robust than individual items from participants (Kline, 2000).

5.6.4.1 Analysis of User Reading Experience scores (UREs) for Font Type, Age Group, and Nationality

An three way ANOVA on the URE scores for the effects of Font Type, Age Group and Nationality found that there was no significant difference due to Age Group ($F(1,68)=.10$, n.s.). However, Font Type and Nationality both had significant main effects (Font Type: $F(1,68)=7.05$, $p<.01$, $\eta_p^2=.09$; Nationality: $F(1,68)=4.38$, $p<.05$, $\eta_p^2=.06$). There was significant interaction between Font Type and Nationality ($F(1,68)=102.15$, $p<.001$, $\eta_p^2=.60$). There were no other interactions between variables.

The main effect for Font Type was that Sans serif font (Mean sans serif=3.33, SD=1.00) were significantly lower in URE scores than serif fonts (Mean serif=3.86, SD=0.81). One sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for serif and san serif were both significantly higher than neutral (serif: $t(71)=9.04$, $p<.001$; sans serif: $t(71)=2.79$, $p<.01$).

The main effect for Age Group was that the UK participants (Mean=3.74, SD=0.79) had URE scores significantly higher than Thai participants (Mean=3.49, SD=1.03). One sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for the UK and Thai participants were both significantly higher than neutral (UK participants: $t(59)=7.15$, $p<.001$; Thai participants: $t(83)=4.38$, $p<.001$).

A Scheffé post hoc analyses was used to test the specific differences in the interaction between Font Type and Nationality. Table 5.3 shows the pattern of the observed t - values for this interaction.

Table 5.3: Observed t - values for all pairs of Font Type and Nationality combination for User Reading Experience scores

		UK		Thai	
		Serif	**Sans serif**	**Serif**	**Sans serif**
UK	Serif	-	-5.37**	-6.13**	-2.70
	Sans serif		-	-5.54**	8.69**
Thai	Serif	-		-	9.47**
	Sans serif	-			-

** p <.01   critical t at 95% confidence level: 3.46, at 99%: 4.59
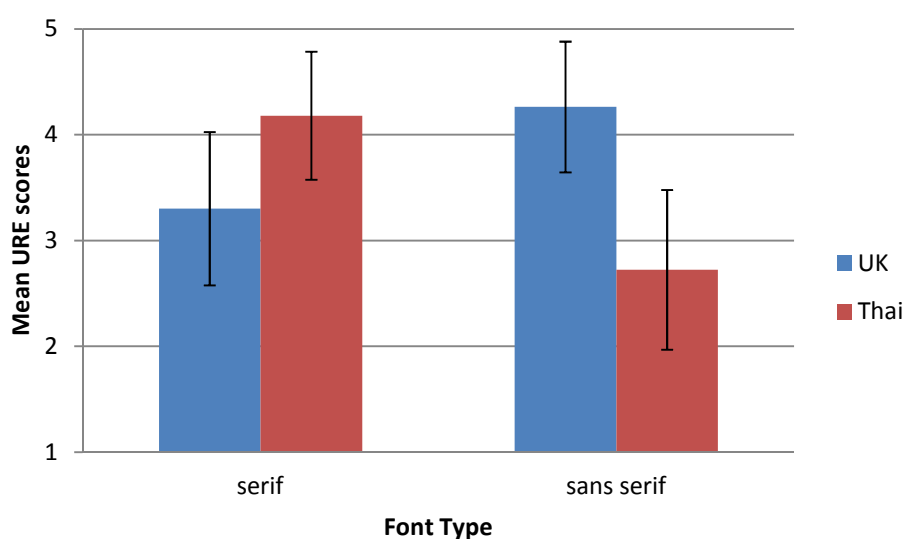
Figure 5.10: Mean URE scores for Font Type for the UK and Thai participants

For UK participants, URE scores for the serif font were significantly lower than the sans serif font (Mean serif=3.30, SD=0.72; Mean sans serif=4.18, SD=0.61, Observed t=-5.37, p<.01, critical t at 95% confidence level: 3.46, at 99%: 4.59). But for Thai participants, URE scores for serif were significantly higher than sans serif (Mean serif=4.26, SD=0.61; Mean sans serif=2.72, SD=0.75, Observed t =9.47, p<.01), This interaction is shown in Figure 5.10.

In addition, the UK participants' URE scores for serif were significantly lower than Thai participants URE for serif (Mean serif for UK participants=3.30, SD=0.72; Mean serif for Thai participants=4.26, SD=0.61, Observed t =-6.13, p<.01), and the UK participants URE scores for sans serif were significantly higher than Thai participants URE for serif (Mean sans serif for UK participants=4.18, SD=0.61; Mean serif for Thai

participants=4.26, SD=0.61, Observed t=-5.54, p<.01), and the UK participants URE scores for sans serif were significantly higher than Thai participants URE for sans serif (Mean sans serif for UK participants=4.18, SD=0.61; Mean sans serif for Thai participants=2.72, SD=0.75, Observed t =8.69, p<.01).

One sample t-tests against the neutral mid-point rating of 3 showed that, for UK participants, the mean ratings for serif and sans serif were both significantly higher than neutral (serif: t(29)=2.27, p<.05; sans serif: t(29)=10.45, p<.001). For Thai participants, the mean ratings for serif was significantly higher than neutral, (t(41)=13.51, p<.001) while the mean ratings for sans serif was significantly lower than neutral, t(41)=-2.38, p<.05).

5.6.4.2 Analysis of User Reading Experience Scores (UREs) for Font Size, Age Group, and Nationality

A three way ANOVA of the effects of Font Size, Age Group and Nationality on User Reding Experience Scores (UREs) found that Font Size and Nationality had significant effects (Font Size: $F(2,136)=200.25$, p<.001, $\eta_p^2=.75$; Nationality: $F(1,68)=7.41$, p< .01, $\eta_p^2=.15$). However there was no significant effect for Age Group ($(1,68)=.47$, n.s.). There were significant interactions between Font Size and Age Group ($F(2,136)=6.32$, p<.01, $\eta_p^2=.09$), and between Font Size, Age Group, and Nationality ($F(2,136)=4.20$, p<.05, $\eta_p^2=.06$). There were no other interaction effects.

A Scheffé post hoc analysis on the specific Font Size differences showed that the mean URE score for 12 point (Mean=2.61 SD=0.69) was significantly lower than the mean URE score for 14 point (Mean=3.81 SD=0.64, Observed ts=-15.25, p<.01, critical t at 95%: 2.47, at 99%: 3.09) or the mean URE score for 16 point (Mean=4.32 SD=0.57, Observed t =-15.63, p<.01), while the mean URE score for 14 point (Mean=3.81 SD=0.64) was significantly lower than the mean URE score for 16 point (Mean=4.32 SD=0.57, Observed t =-5.67, p<.01), This interaction is shown in Figure 5.11.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean URE for 12 point was significantly lower than neutral (t(71)=-4.77, p<.001), but the mean URE for 14 point and 16 point were both significantly higher than neutral (14 point: t(71)=10.76, p<.001; 16 point: t(71)=19.79, p<.001).
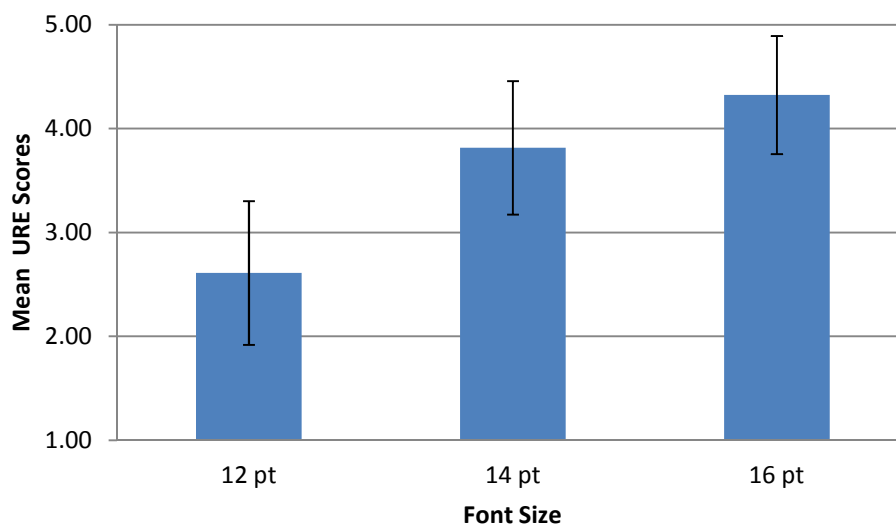
Figure 5.11: Mean URE scores for the different Font Sizes

The main effect for Natioality was that Thai participants were significantly more positive in their URE scores than the UK participants (Mean Thai participants=3.72 SD=0.84; Mean UK participants=3.39 SD=1.08). One sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for the UK and Thai participants were both significantly higher than neutral (UK participants: t (89) =3.40, p < .01; Thai participants: t (125) = 9.65, p < .001).

A Scheffé post hoc analysis was used to test the specific differences in the interaction between Font Size and Age Group, and the interaction between Font Size, Age Group and Nationality. Table 5.4 shows the pattern of observed t-values for the Scheffé post hoc analysis of the two way interaction between Font Size and Age Group. For Younger Adults, 12 point (Mean 12 point=2.66 SD=0.74) scored significantly lower than 14 (Mean 14 point=3.95 SD=0.61) (Observed t=-10.86, p<.01, critical t at 95%: 3.91, at 99%: 4.88) or 16 point (Mean 16 point=4.18 SD=0.61) (Observed t= -8.83, p<.01). But there was no difference in rating between 14 and 16 point. For Older Adults, 12 point (Mean 12 point=2.56 SD=0.64) scored significantly lower than 14 point (Mean 14 point=3.66 SD=0.65) (Observed t=-11.13, p<.01) or 16 point (Mean 16 point= 4.49 SD=0.46) (Observed t=-16.65, p<.01) while 14 point scored significantly lower than 16 point (Observed t=-8.70, p<.01). This interaction is shown in Figure 5.12.

In addition there were significant differences in URE scores between Younger Adults 12 point and Older Adults 14 point, between Younger Adults 12 point and Older Adults 16 point, between Younger Adults 14 point and Older Adults 12 point, between

Younger Adults 14 point and Older Adults 16 point, and between Younger Adults 16 point and Older Adults 12 point

One sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for the 12 point size for both the Younger and Older Adults were significantly lower than neutral (12 point Younger Adults: t(38)=-2.89, p<.01; 12 point Older Adults: t(32)=-4.00, p<.001). While all the other mean URE scores were significantly higher than neutral (14 point Younger Adults: t(38)=-9.64, p<.01; 14 point Older Adults: t(32)=5.82, p<.001; 16 point Younger Adults: t(38)=-12.06, p<.01; 16 point Older Adults: t(32)=18.45, p<.001).

Table 5.4: Observed t-values between all pairs of Font Size and Age Group combination for User Reading Experience scores

		Younger Adults			Older Adults		
		12 pt	**14 pt**	**16 pt**	**12 pt**	**14 pt**	**16 pt**
Younger Adults	12 pt	-	-10.86**	-8.83**	0.62	-5.34**	-11.87**
	14 pt		-	-1.78	8.35**	1.96	-4.24*
	16 pt			-	10.27**	3.43	-2.48
Older Adults	12 pt				-	-11.13**	-16.65**
	14 pt					-	-8.70**
	16 pt						-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.91, at 99%: 4.88

Figure 5.12: Mean URE scores for different Font Sizes for Younger and Older Adults

A Scheffé post hoc analysis was used to investigate the specific differences in the interaction between Font Size, Age Group, and Nationality. This interaction is shown in Figure 5.13. Table 5.5 shows the pattern of observed t-values for the Scheffé post hoc analysis of the two way interaction between Font Size and Age Group for the UK participants. For UK Younger Adults, 12 point scored significantly lower than 14 point (Mean 14 point=3.93 SD=0.77) (Observed t-values=-13.25, p<.01, critical t at 95%: 5.80, at 99%: 7.24) and 16 point (Mean 16 point=4.17 SD=0.64) (Observed t-values=-7.36, p<.01). But there was no difference in rating between 14 and 16 point. For UK Older Adults, 16 point (Mean 16 point=4.28 SD=0.45) was scored significantly higher than 12 point (Mean 12 point=2.33 SD=0.72) (Observed t-values=-8.15, p<.01) and 14 point (Mean 14 point=3.33 SD=0.59) (Observed t-values=-7.34, p<.01). But there was no significant difference between 12 and 14 point.

In addition, URE scores for 12 point for UK Younger participants were significantly lower than 16 point for UK Older participants. URE scores for 16 point for UK Younger participants were significantly higher than URE scores for 12 point for UK Older participants.

**Table 5.5: Observed t-values between all pairs of Font Size, Age Group
combination for User Reading Experience scores for the UK participants**

		Younger Adults			Older Adults		
		12 pt	**14 pt**	**16 pt**	**12 pt**	**14 pt**	**16 pt**
Younger Adults	12 pt	-	-13.25**	-7.36*	-0.47	-4.35	-8.50**
	14 pt		-	-1.03	5.67	2.26	-1.42
	16 pt			-	7.30**	3.61	-0.52
Older Adults	12 pt				-	-5.45	-8.15**
	14 pt					-	-7.34**
	16 pt						-

* p<.05, ** p <.01   critical t at 95% confidence level: 5.80, at 99%: 7.24

Table 5.6 show the pattern of observed t-values for the Scheffé post hoc analysis of the two way interaction between Font Size and Age Group for the Thai participants. For Thai Younger Adults, 12 point (Mean 12 point=3.05 SD=0.45) scored significantly lower than 14 point (Mean 14 point=3.97 SD=0.46) (Observed t=-6.11, p<.05, critical t at 95%: 5.80, at 99%: 7.24) and 16 point (Mean 16 point=4.19 SD=0.60) (Observed t=-5.91, p<.05). But there was no significant difference in rating between 14 and 16 point. For Thai Older Adults, 12 point (Mean 12 point=2.68 SD=0.56) scored significantly lower than 14 point (Mean 14 point=3.84 SD=0.62) (Observed t=-9.95, p<.01) and 16 point. (Mean 16 point=4.62 SD=0.62) (Observed t =-15.21, p<.01). In addition, 14 point scored significantly lower than 16 point (Observed t=-5.86, p<.05).

In addition, URE scores for 12 point for Thai Younger participants werer significantly lower than for 16 point for Thai Older participants. URE scores for 12 point for Thai Older participants were significantly than for both 14 point for Thai Younger participants and URE scores for 16 point for Thai Younger participants.

Table 5.6: Observed t-values between all pairs of Font Size, Age Group combination for User Reading Experience scores for Thai participants

		Younger Adults			Older Adults		
		12 pt	14 pt	16 pt	12 pt	14 pt	16 pt
Younger Adults	12 pt	-	-6.11*	-5.91*	2.32	2.32	-11.45**
	14 pt		-	-1.60	8.12**	0.76	-4.70
	16 pt			-	8.39**	1.85	-2.64
Older Adults	12 pt				-	-9.95**	-15.21**
	14 pt					-	-5.86*
	16 pt						-

* p<.05, ** p <.01   critical t at 95% confidence level: 5.80, at 99%: 7.24
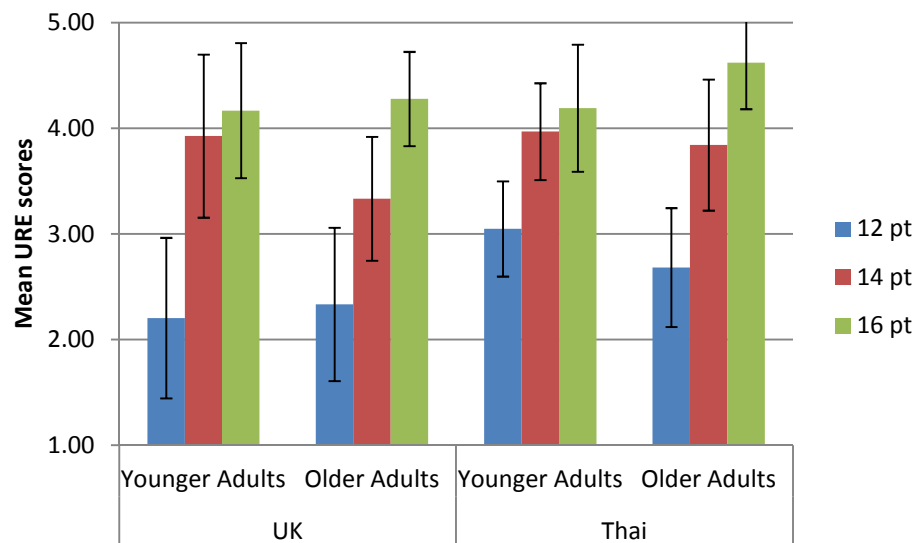
Figure 5.13: Mean URE scores for different Font Sizes for Younger and Older Adults for the UK and Thai participants

For UK participants, one sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for the 12 point size for both the Younger and Older Adults were significantly lower than neutral (12 point Younger Adults: $t(17)=-4.45$, $p<.001$; 12 point Older Adults: $t(11)=-3.18$, $p<.01$). While all the other mean scores were significantly higher than neutral (14 point Younger Adults: $t(17)=-5.09$, $p<.001$; 16 point Younger Adults: $t(17)=7.75$, $p<.001$; 16 point Older Adults: $t(11)= 9.90$, $p<.001$), except

14 point for Older Adults which was rated not significantly different from the neutral mid-point (t(32)=5.82, n.s.).

For Thai participants, one sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for the 12 point size for the Younger not significantly different from the neutral mid-point (t(20)=0.48, n.s.). The mean URE scores for 12 point size for Older Adults were significantly lower than neutral (t(20)=-2.59, p<.05). While all the other mean URE scores were significantly higher than neutral (14 point Younger Adults: t(20)=-9.69, p<.001; 16 point Younger Adults: t(20)=9.06, p<.001; 14 point Older Adults: t(20)=6.21, p<.001; 16 point Older Adults: t(20)=16.92, p<.001).

5.6.4.3 Participants' overall preference ratings on combinations of Font Type and Font Size

A four way ANOVA on overall preference ratings on combination of Font Type and Font Size found that Font Type, Font Size, and Nationality all had significant main effects (Font Type: $F(1,68)=7.20$, p<.01, η_p^2 =.10; Font Size: $F(2,136)=133.98$, p<.001, η_p^2 =.66; Nationality: $F(1, 86)=4.26$, p<.05, η_p^2 =.06) while Age Group did not have a significant effect ($F(1,68)=0.15$, n.s.). There were significant interactions between Font Type and Nationality ($F(1,68)=158.98$, p<.001, η_p^2 =.70), between Font Size and Age Group ($F(2,136)=8.25$, p<.001, η_p^2 =.11), and between Font Type, Font Size, Age Group, and Nationality ($F(2,136)=5.42$, p<.01, η_p^2 =.07).

The main effect for Font Type was that overall rating for serif (Mean serif=3.29, SD=1.26) was significantly higher than for sans serif (Mean sans serif=2.80, SD=1.34).

One sample t-tests against the neutral mid-point rating of 3 showed that the mean ratings for serif was significantly higher than neutral (t(215)=3.40, p<.01) but sans serif was significantly lower than neutral (t(215)=-2.18, p<.05).

A Scheffé post hoc analysis found that overall preference ratings for 12 point (Mean 12 point= 20.9, SD = 1.15) were significantly lower than 14 point (Mean 14 point=3.30 SD=1.13) (Observed t=-15.18, p<.01, critical t at 95%: 2.47, at 99%: 3.09) and 16 point (Mean 16 point=3.75, SD=1.10) (Observed t=-13.15, p<.01) while overall preference ratings for 14 point (Mean 14 point=3.30 SD=1.13) were significantly lower than 16 point (Mean 16 point=3.75, SD=1.10) (Observed t=-3.65, p<.01), This interaction is shown in Figure 5.14.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean rating for 12 point was significantly lower than neutral (t(143)=9.53, p<.001), but the mean ratings for 14 point and 16 point were both significantly higher than neutral (14 point: t(143)=3.18, p<.01; 14 point: t(143)=8.18, p<.001).
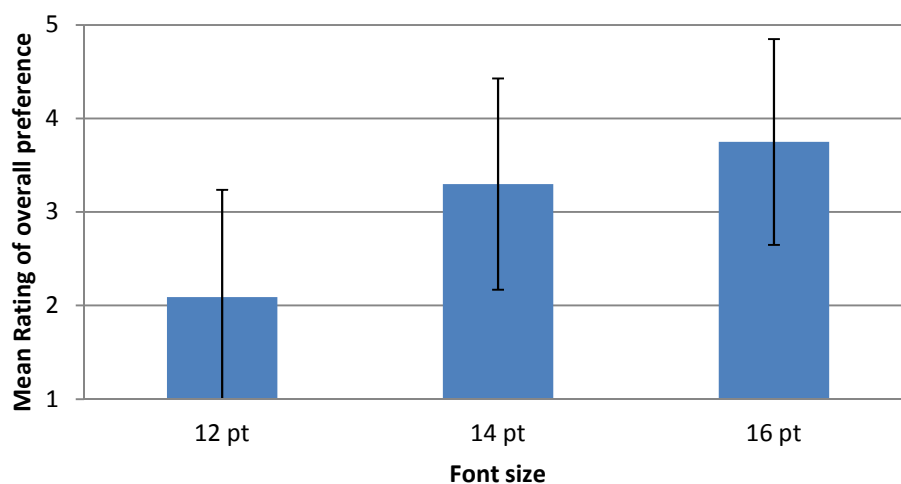
Figure 5.14: Mean Rating of overall preference for different Font Sizes

The main effect for Nationality was that the UK participants were significantly more positive in their overall ratings of preference than the Thai participants, (F(1,68)=4.26, p<.05) (UK participants: Mean=3.15 SD=1.32; Thai participants: Mean=2.97 SD=1.32).

One sample t-tests against the neutral mid-point rating of 3 showed that the mean rating for both the UK participants and Thai participants were not significantly different from neutral (UK participants: t(179)=1.52, n.s.; Thai participants: t(251)=-3.33, n.s.).

A Scheffé post hoc analysis was used to test the specific differences in the two way interaction between Font Type and Nationality. Table 5.7 shows the observed t-values for each combination of Font Type and Nationality. For UK participants, serif was rated significantly lower in overall preference than sans serif (Mean serif=2.66, SD=1.23; Mean sans serif=3.54, SD=1.24, Observed t=-8.40, p<.01, critical t at 95%: 3.46, at 99%: 4.59). But for Thai participants, serif was rated significantly higher in preference than sans serif (Mean serif=3.75, SD=1.08; Mean sans serif=2.20, SD=1.07, Observed t=13.63, p<.01). This interaction is shown in Figure 5.15.

In addition, the mean rating of UK participants for serif (Mean UK participants for serif=2.66, SD=1.23) was significantly lower than Thai participants for serif (Mean Thai participants for serif=3.75, SD=1.08, Observed t=-6.91, p<.01), and UK participants rating for sans serif (Mean UK participants for sans serif=3.54, SD=1.24) was

significantly higher than Thai participants ratings (Mean UK participants for sans serif=2.20, SD=1.07, Observed t=9.22, p<.01). There were no other significant differences.

One sample t-tests against the neutral mid-point rating of 3 showed that, for UK participants, the mean rating for serif was significantly lower than neutral (serif: t(89)=-2.66, p<.001) while the mean ratings for sans serif was significantly higher than neutral (sans serif: t(89)=4.97, p<.001). For Thai participants, the mean ratings for serif was significantly higher than neutral, (t(125)=7.75, p<.001) while the mean ratings for sans serif was significantly lower than neutral, t(125)=-8.44, p<.001).

Table 5.7: Observed t-values between all pairs of Font Type and Nationality combination for Participants' overall preference

		UK		Thai	
		Serif	**Sans serif**	**Serif**	**Sans serif**
UK	Serif	-	-8.40**	-6.91**	3.01
	Sans serif		-	-0.13	9.22**
Thai	Serif	-	-	-	13.63**
	Sans serif	-	-	-	-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.46, at 99%: 4.59
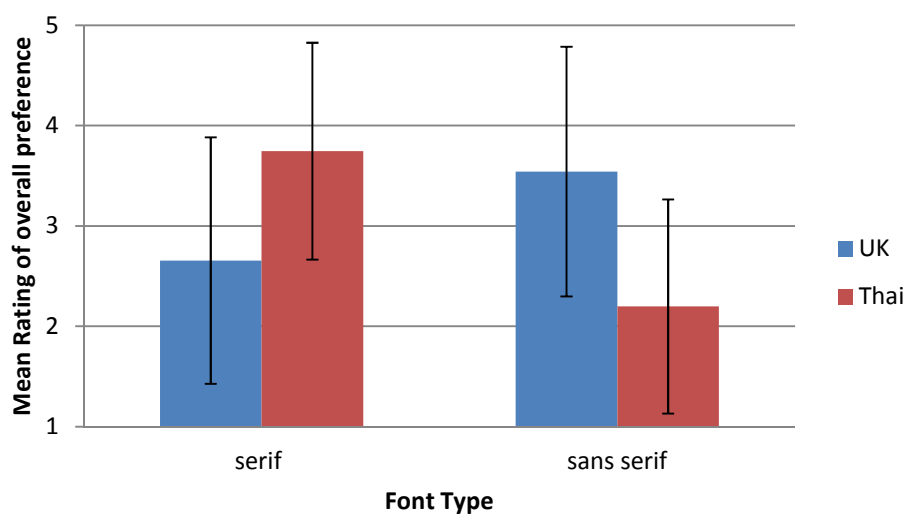
Figure 5.15: Mean Rating of overall preference for serif and sans serif for UK and Thai participants

A Scheffé post hoc analysis was used to test the specific differences in the two way interaction between, Table 5.8 shows that the observed t-values for each combination of Font Size and Age Group. For Younger Adults, 12 point (Mean 12 point=2.19, SD=1.28) was rated significantly lower than 14 point (Mean 14 point=3.47, SD=1.14) (Observed t-values=-10.37, p<.01, critical t at 95%: 3.91, at 99%: 4.88) or 16 point (Mean 16 point=3.58, SD=1.05) (Observed t-values=-7.50, p<.01) while 14 point and 16 point were not significantly different from each other.

For Older Adults, 12 point (Mean 12 point=1.97, SD=0.96) was rated significantly lower than both 14 point (Mean 14 point=3.09, SD=1.09) (Observed t-values=-12.78, p< 0.01) and 16 point (Mean 16 point=3.95, SD=1.13) (Observed t-values=-16.93, p<.01) while 14 point was rated significantly lower than 16 point (Observed t-values=-6.59, p<.01). This interaction is shown in Figure 5.16.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean ratings for the 12 point size for both the Younger and Older Adults were significantly lower than neutral (12 point Younger Adults: t(77)=-5.57, p<.001; 12 point Older Adults: t (65)=-8.72, p<.001). While the other ratings were significantly higher than neutral (14 point Younger Adults: t(77)=-3.69, p<.001; 16 point Younger Adults: t(77)=4.85, p<.001; 16 point Older Adults: t(65)=6.87, p<.001), except 14 point for Older Adults which was not significant different from neutral, (t(65)=0.68, n.s.).

Table 5.8: Observed t-values between all pairs of Font Size and Age Group combination for Participants' overall preference

		Younger Adults			Older Adults		
		12 pt	**14 pt**	**16 pt**	**12 pt**	**14 pt**	**16 pt**
Younger Adults	12 pt	-	-10.37**	-7.50**	1.16	-4.49**	-8.69**
	14 pt		-	-0.66	8.49**	2.05	-2.36
	16 pt			-	9.51**	2.72	-2.07
Older Adults	12 pt				-	-12.78**	-16.93**
	14 pt					-	-6.59**
	16 pt						-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.91, at 99%: 4.88

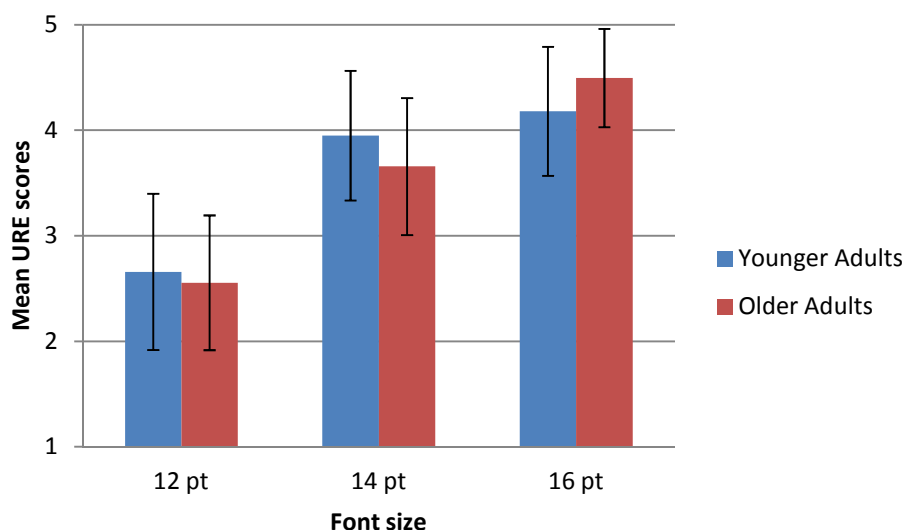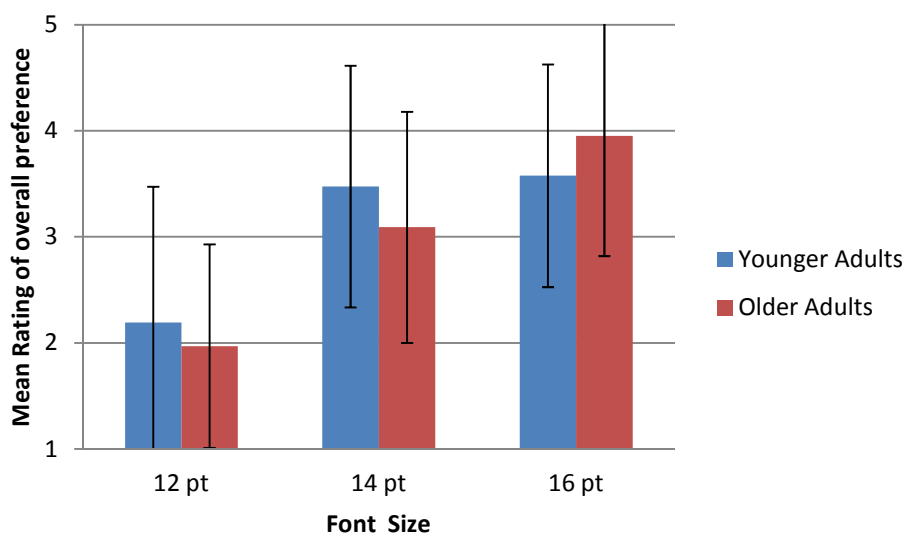**Figure 5.16: Mean Rating of overall preference for different Font Size for Younger and Older Adults**

A Scheffé post hoc analysis was used to test the specific differences in the four way interaction between Font Size, Font Type, Age Group and Nationality. In order to provide information for recommendation on Font Type and Font Size for both the UK and Thai and for both the Younger and Older Adults, the analysis was separated to look at the interaction between Font Type and Font Size for each of four groups of participants: UK Younger Adults, UK Older Adults, Thai Younger Adults, and Thai Older Adults. In these cases, the critical t – Scheffe value was 3.91 for the 95 per cent confidence interval, and 4.88 for the 99 per cent confidence interval.

Table 5.9 shows that the observed t-values between each combination of Font Type and Font Size for the UK Younger Adults. Serif 12 point (Mean=1.44, SD=0.86) was rated significantly lower than both serif 14 point (Mean=3.11 SD=0.83) (Observed t=-7.29, p<.01, critical t at 95%: 3.91, at 99%: 4.88) and serif 16 point (Mean =3.39, SD=0.85) (Observed t=-7.10, p<.01), while serif 14 point and serif 16 point were not significantly different from each other (Observed t=-1.10, n.s.). In addition, sans serif 12 point was rated significantly lower than sans serif 14 point (mean=4.11 SD=0.90) (Observed t=-7.16, p<.01). Sans serif 12 point and 16 point and sans serif 14 point and16 point were not significantly different from each other. This interaction is shown in Figure 5.17.

In addition, serif 12 point was rated significantly lower than both sans serif 14 point and 16 point.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean ratings for the serif 12 point was significantly lower than neutral, (t(17)=-7.71, p<.001). While sans serif 14 point, and 16 point were significantly higher than neutral (sans serif 14 point: t(17)=5.24, p<.001; sans serif 16 point: t(17)=4.03, p<.01). In addition, serif 14 point, serif 16 point, and sans serif 12 point were not significantly different from neutral (serif 14 point: t(17)=0.57, p>0.05; serif 16 point: t(17) =1.94, p>.005, sans serif 12 point: t(17)=-1.64, p>.05).

Table 5.9: Observed t-values between all pairs of Font Types and Font Sizes combination for Participants' overall preference for the UK Younger Adults

		Serif			Sans serif		
		12 pt	14 pt	16 pt	12 pt	14 pt	16 pt
serif	12 pt	-	-7.29**	-7.10**	-3.83	-8.00**	-6.20**
	14 pt		-	-1.10	1.49	-3.09	-2.41
	16 pt			-	2.01	-1.91	-2.61
Sans serif	12 pt				-	-7.16**	-3.17
	14 pt					-	0.15
	16 pt						-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.91, at 99%: 4.88
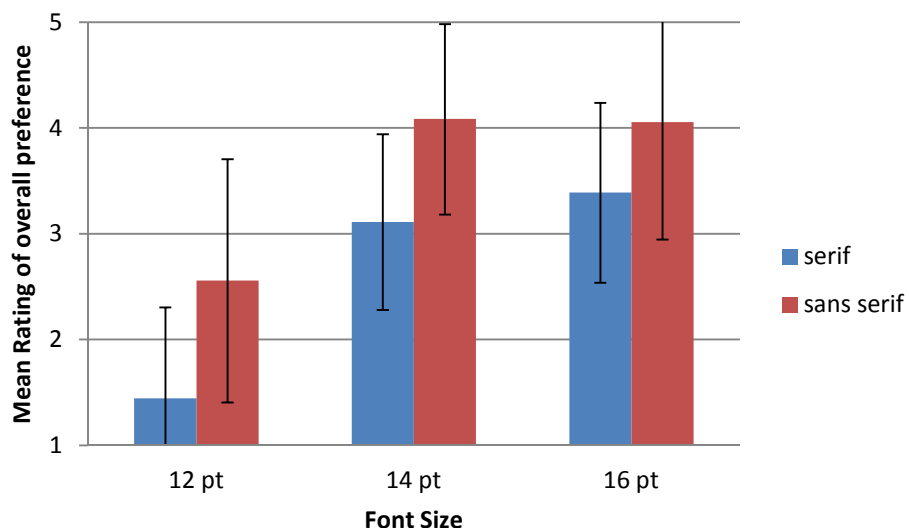
Figure 5.17: Mean Rating of overall preference for serif and sans serif font for different Font Sizes for the UK Younger Adults

Table 5.10 shows that the observed t-values between each combination of Font Type and Font Size for the UK Older Adults. Serif 12 point (Mean=1.50, SD=0.67) was rated significantly lower than both serif 14 point (Mean=2.67 SD=0.89) (Observed t=-7.00, p<.01, critical t at 95%: 3.91, at 99%: 4.88) and 16 point (Mean=3.83, SD=1.03) (Observed t=-10.38, p<.01). Serif 14 point (Mean=2.67 SD=0.89) was rated significantly lower than 16 point (Mean=3.83, SD=1.03) (Observed t=-4.31, p<.05).

In addition, sans serif 12 point (Mean=2.75, SD=0.75) was rated significantly lower than both 14 point (Mean=3.58 SD=1.08) (Observed t=-4.02, p<.05) and 16 point (Mean=4.92, SD=0.29) (Observed t=-8.99, p<.01). However, between sans serif 14 point and 16 point were not significantly different from each other.

Finally, serif 12 point was rated significantly lower than sans serif 12 point, sans serif 14 point, and sans serif 16 point. In addition, serif 14 point was rated significantly lower than sans serif 16 point. This interaction is shown in Figure 5.18.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean ratings for the serif 12 point was significantly lower than neutral, (t(11)=-7.71, p<.001). Serif 16 point and sans serif 16 point were significantly higher than neutral (serif 16 point: t(11)=2.80, p<.05; sans serif 16 point: t(11)=23.00, p<.001). In addition, serif 14 point, sans serif 12 and 14 point were not significantly difference from neutral (serif 14 point: t(11)=-1.30, n.s.; sans serif 12 point: t(11)=1.15, n.s., sans serif 14 point: t(11)=-1.86, n.s.).

Table 5.10: Observed t-values between all pairs of Font Types and Font Sizes combination for Participants' overall preference for the UK Older Adults

		Serif			Sans serif		
		12 pt	**14 pt**	**16 pt**	**12 pt**	**14 pt**	**16 pt**
serif	12 pt	-	-7.00**	-10.38**	-5.00**	-6.60**	-17.70**
	14 pt		-	-4.31*	-0.36	-3.19	-9.00**
	16 pt			-	3.23	0.39	-3.77
Sans serif	12 pt				-	-4.02*	-8.99**
	14 pt					-	-3.75
	16 pt						-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.91, at 99%: 4.88
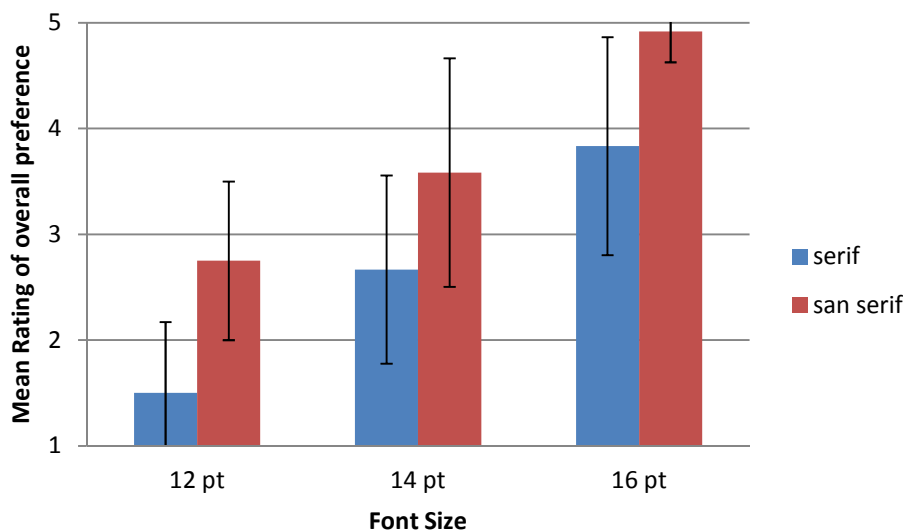
Figure 5.18: Mean Rating of overall preference for serif and sans serif font for different Font Size for the UK Older Adults

Table 5.11 shows that the observed t-values between each combination of Font Type and Font Size for Thai Younger Adults. Serif 12 point (Mean=3.48, SD=0.98) serif 14 point (Mean=4.14 SD=0.85) serif 16 point (Mean=3.81, SD=1.08) were not significantly difference from each others (Observed t between serif 12 point and serif 14 point=-2.87, n.s.; Observed t between serif 12 point and serif 16 point=-0.87, n.s.; Observed t between serif 14 point and serif 16 point=-0.98, n.s., critical t at 95%: 3.91, at 99%: 4.88).

In addition, sans serif 12 point (Mean=1.24, SD=0.54) was rated significantly lower than both sans serif 14 point (Mean=2.57 SD=1.08) (Observed t=-5.29, p<.01), and 16 point (Mean=3.10, SD=0.94) (Observed t=-8.83, p<.01). However, between sans serif 14 point and 16 point were not significantly different from each other.

Finally, sans serif 12 point was rated significantly lower than any serif size. In addition, sans serif 14 point was rated significantly lower than serif 14 point This interaction is shown in Figure 5.19.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean ratings for sans serif 12 point was significantly lower than neutral, (t(20)=-14.98, p<.001). While all serif sizes were significantly higher than neutral (serif 12 point: t(20)=2.23, p<.05; serif 14 point: t(20)=6.14, p<.001; serif 16 point: t(20)=3.44, p<.05). In addition, sans serif 14 point and 16 point were not significantly different from neutral (sans serif 14 point: t(20)=-1.83, n.s.; sans serif 16 point: t(20)=0.46, n.s.).

Table 5.11: Observed t-values between all pairs of Font Types and Font Sizes combination for Participants' overall preference for Thai Younger Adults

		Serif			Sans serif		
		12 pt	**14 pt**	**16 pt**	**12 pt**	**14 pt**	**16 pt**
serif	12 pt	-	-2.87	-0.87	8.70**	2.80	1.09
	14 pt		-	0.98	12.75**	4.59*	3.13
	16 pt			-	11.46**	3.46	2.37
Sans serif	12 pt				-	-5.29**	-8.83**
	14 pt					-	-2.33
	16 pt						-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.91, at 99%: 4.8
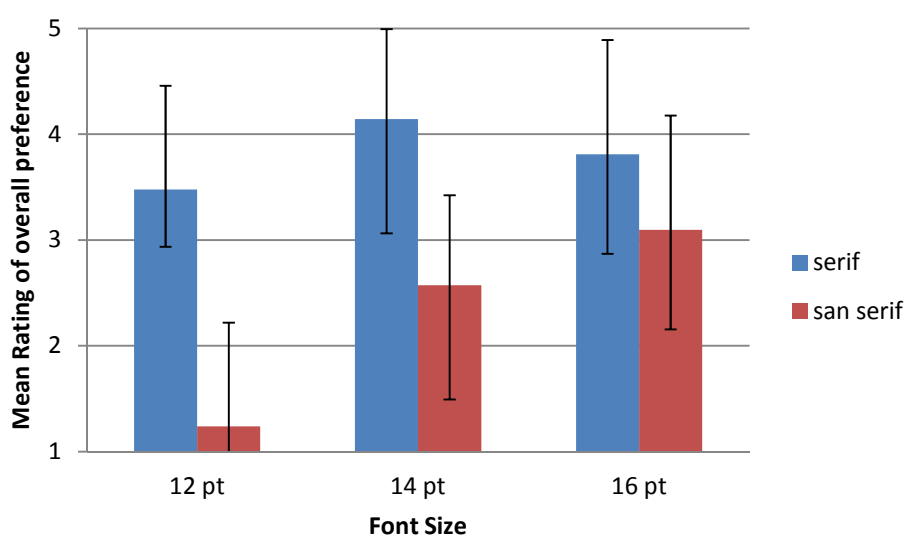
Figure 5.19: Mean Rating of overall preference for serif and sans serif font for different Font Size for Thai Younger Adults

Table 5.12 shows that the observed t-values between each combination of Font Type and Font Size for Thai Older Adults. Serif 12 point (Mean=2.57, SD=0.87) was rated significantly lower than both serif 14 and point (Mean=3.86 SD=0.79) (Observed t=-8.22, p<.01, critical t at 95%: 3.91, at 99%: 4.88), and serif 16 point (Mean=4.62, SD=0.74) (Observed t=-7.54, p<.01) while serif 14 point and serif 16 point were not significantly different from each other.

In addition, sans serif 12 point (Mean=1.19, SD=0.40) was rated significantly lower than both sans serif 14 point (Mean=2.29 SD=0.78) (Observed t= -6.53, p<.01), and

sans serif 16 point (Mean=2.81 SD=0.75) (Observed t =-11.09, p<.01). Sans serif 14 and16 point were not significantly different from each other. This interaction is shown in Figure 5.20.

Finally, sans serif 12 point was rated significantly lower than all serif sizes. In addition, sans serif 14 point was rated significantly lower than serif 14 or 16 point. Sans serif 16 point was rated significantly lower than serif 14 point and 16 point.

Table 5.12: Observed t-values between all pairs of Font Types and Font Sizes combination for Participants' overall preference for Thai Older Adults

		Serif			Sans serif		
		12	**14**	**16**	**12**	**14**	**16**
serif	12 pt	-	-8.22**	-7.54**	7.32**	1.55	-1.16
	14 pt		-	-3.07	15.36**	7.36**	5.55**
	16 pt			-	21.05**	8.64**	7.11**
Sans serif	12 pt				-	-6.53**	-11.09**
	14 pt					-	-2.75
	16 pt						-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.91, at 99%: 4.88
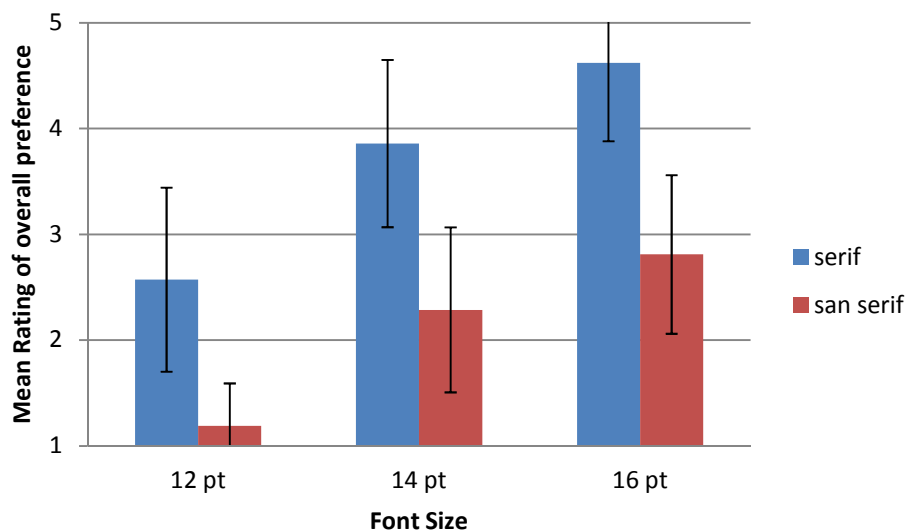
Figure 5.20: Mean Rating of overall preference for serif and sans serif font for different Font Size for Thai Older Adults

One sample t-tests against the neutral mid-point rating of 3 showed that the mean ratings for the sans serif 12 and 14 point, and serif 12 point was significantly lower than neutral, (sans serif 12 point: $t(20)=-20.61$, $p<.001$; sans serif 14 point: $t(20)=-4.18$, $p<.001$; serif 12 point: $t(20)=-2.26$, $p<.05$). Serif 14 and 16 point were significantly higher than neutral (serif 14 point: $t(20)=4.95$, $p<.001$; serif 16 point: $t(20)=10.03$, $p<.001$). Finally, sans serif 16 point was not significantly different from neutral (sans serif 16 point: $t(20)=-1.16$, n.s.).

5.6.5 Predicting reading performance from Attitudes to the Web Scale (ATWS)

Linear Regression was computed to predict participants' reading performance as measured by Time spent per web page from the three Factors of the Attitudes to the Web Scale (ATWS) , Age Group, and Nationality.

The linear regression produced a significant overall proportion of the variance predicted ($F(5,71)=7.60$, $p<.01$, Adjusted $R^2=0.32$). The Confidence Factor of the ATWS and Nationality were significant individual predictors (Confidence Factor: $t=-2.49$, $p<.05$; Nationality: $t=4.21$, $p<.001$). However, the Performance Factor, Fashion Factor, and Age Group were not significant predictors. The results are summarized in Table 5.13.

Table 5.13 The B-values, t-values, and significance levels for the linear regression predicting time per web page from the ATWS factors, Age Group, and Nationality

	B	**t**	**Sig.**
Confidence Factor	-14.01	-2.49	.02
Performance Factor	14.30	1.93	.06
Fashion Factor	-4.96	-1.08	.28
Age Group	1.99	0.33	.74
Nationality	24.80	4.21	.00

Figure 5.21 shows a scatterplot of the correlation between Confidence Factor and Reading time per web page. The higher rating on Confidence Factor, the less Reading time per web page.
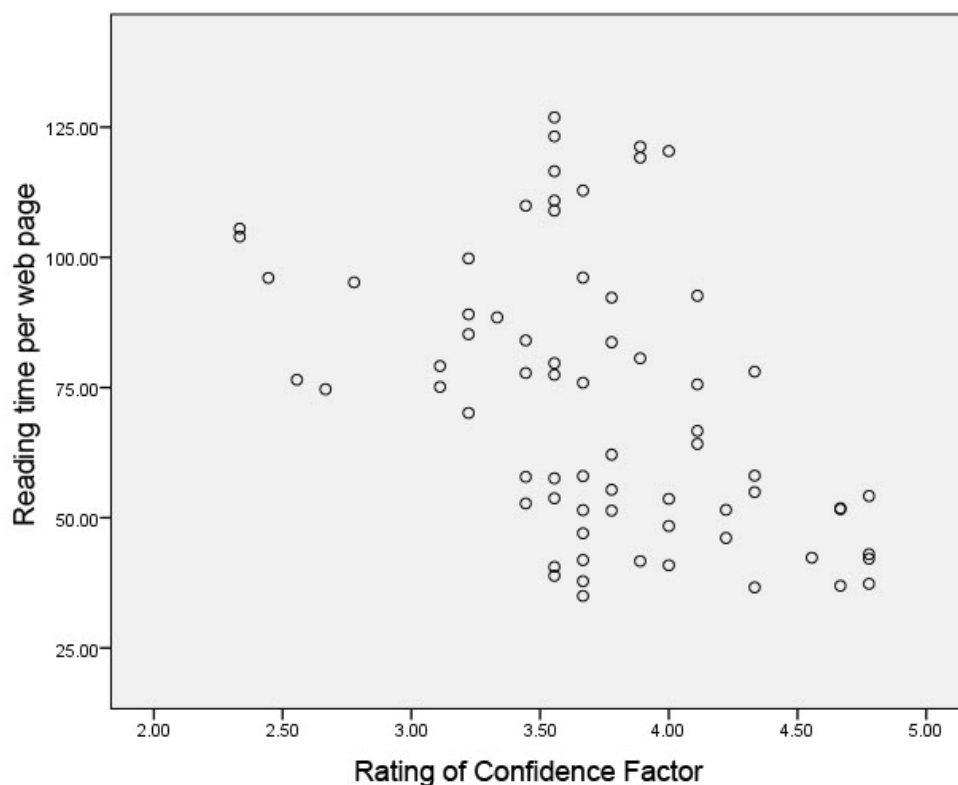
Figure 5.21: A scatterplot of the correlation between Confidence Factor and Reading time per web page.

For the Nationality predictor, Thai participants spent longer Reading time per web page than the UK participants (Mean Thai participants=82.82 sec SD=28.37, Mean UK participants=55.78 sec SD=18.93). This difference is in line with the findings .

5.7 Discussion

This study investigated the effect of Font Type and Font Size on performance and preference measures of skim reading text from a computer screen for Younger and Older Adults in the UK and Thailand.

On the performance measures, the first thing to consider is whether participants were changing their skim reading behaviour using a speed-accuracy trade-off. To investigate this, I asked participants four multiple choice questions about each page they skim read. There were no significant differences in the accuracy of their answers

due to Font Type or Size or Age Group, although there was a small effect size[2] significant effect of Nationality. UK participants were more accurate than Thai Participants. This may have been because UK participants would have been more familiar with material about the Olympic Games than the Thai participants and found it easier to understand and assimilate new material. Most importantly, there was no correlation between the time spent per page and the percentage of correct answers, so there was no tendency for participants to spend longer and be more accurate. Thus it is reasonable to conclude that there was not speed-accuracy trade-off.

Turning to the results on each of the independent variables, for Font Type, on the performance measures, had a medium size and significant overall effect, with participants spending less time skim reading serif fonts than sans serif fonts. There was also a significant interaction between Font Type and Nationality, but the differences within each national group in the skim reading time for Serif versus Sans Serif were not significant in post hoc tests.

On the preference measures, Font Type had a small but significant effect on URE scores, with Serif fonts preferred to Sans Serif fonts. A similar result was obtained from the overall preference rating, with a small sized effect in favour of Serif fonts. This is in line with the performance measures, in that Serif fonts produced shorter times on webpage. However, the URE scores for both Serif and Sans Serif fonts were significantly above the midpoint of the rating scale, showing that either type of font is acceptable to participants.

However, the main effect and the small significant effect of Nationality on URE scores, obscure an interesting medium interaction effect between Font Type and Nationality. UK participants preferred the Latin Sans Serif font significantly more than the Thai participants preferred the Thai equivalent font, whereas Thai participants preferred the Thai Serif font equivalent significantly more than the UK participants. Moreover, UK participants preferred Latin San Serif significantly more than Latin Serif but the URE scores for both font types were significantly above the midpoint of the rating scale. Thai participants preferred Thai Serif font equivalent significantly more than Thai San Serif font equivalent. The URE scores for Thai Serif font equivalent was significantly above the midpoint but for Thai San Serif font equivalent was significantly lower than midpoint.

---

[2] A small effect is one that captures about 1 percent of the variance. In term of standardised difference, a small effect has value approximate 0.25. A medium effect captures about 6 percent of the variability, value is approximate 0.5. A large effect captures at least 15 percent of variability, value is approximate 0.8 (Keppel, 2004).

On overall preference rating, there were medium interaction effect between Font Type and Nationality. Again, UK participants preferred the Latin Sans Serif font significantly more than the Thai participants preferred the Thai equivalent font, whereas Thai participants preferred the Thai Serif font equivalent significantly more than the UK participants. UK participants significantly preferred Latin San Serif more than Latin Serif. The URE scores for Latin San Serif was significantly above the midpoint of the rating scale but the URE scores for Latin Serif was significantly lower than midpoint. Thai participants preferred Thai Serif font equivalent significantly more than Thai San Serif font equivalent. The URE scores for Thai Serif font equivalent was significantly above the midpoint but for Thai San Serif font equivalent was significantly lower than midpoint.

Comparing with previous research, Bernard, Liao and Mills (2001a, 2001b) did a similar study using a detailed reading task. In their work, there was no significant effect on mean time to read reported regarding Font Type (Serif vs. Sans Serif). In current study, Serif fonts was read significantly faster than Sans Serif fonts.

The results of the current study indicate that there are substantial differences between the UK and Thailand in regards to what type of text is best for reading on the web. The current study found that the UK participants preferred the San Serif font while the Thai participants preferred Thai conservative font type which closely related to the Serif font type in Latin alphabet. This might because Thai people are more familiar with Thai conservative font type more than Thai modern font type (closely related to San Serif). The Thai conservative font is usually used in book, newspaper, other print media, and the web content. While Thai modern font type is usually used in advertising (print media, television, online), and some magazines which have younger adults as target audiences. For mobile devices, the good example was the font type which is used in the Apple iOS. Apple used Thai modern font type at the starting of introducing iOS7 but they have changed to Thai conservative font type a few months later. For now, the Thai conservative font type is using for iOS8.

Turning to consider the results for Font Size, font size had no significant effect on skim reading time, and the accuracy of participants answers. However, there are interesting results on overall fatigue and preference measures, URE scores, and overall preference. On the overall fatigue, Font Size had small significant main effect. On URE scores, participants rated overall fatigue for 12 point significantly over than 16 point but there were no different between 12 point and 14 point, and between 14 point and 16 point. Moreover, Font size had a large main effect on URE score. participants least

preferred 12 point but preferred 16 point the most. The URE scores for both 14 point and 16 point were significantly above the midpoint of the rating scale while 12 point was significantly lower than midpoint,

On URE scores, there were small sized interaction effect between Font Size and Age Group, and between Font Size, Age Group, and Nationality. On the interaction between Font Size and Age Group, there was no different on URE scores for each Font Size between Younger and Older Adults. 14 point and 16 point were rated above the midpoint but 12 point was rated lower than midpoint. On the interaction between Font Size, Age Group, and Nationality, UK and Thai Younger Adults preferred 14 point and 16 point over 12 point. Whereas UK Older Adults significantly preferred 16 point over 14 point and 12 point. In addition, Thai Older Adults preferred 16 point the most and preferred 12 point the least. Once comparing the URE scores with the midpoint of the rating scale, 14 point and 16 point were rated above the midpoint by all group of participants while 12 point was rated lower than the mid point by all group of participants, except Thai Younger Adults who rated 12 point around the midpoint of the rating scale.

On overall preference, there was a medium sized main effect of Font Size. 12 point was least preferred while 16 point was the most preferred. 12 point was rated lower than the midpoint but 14 point and 16 point was rated above the midpoint of the rating scale. Moreover, there were small sized interaction effect between Font Size and Age Group. For Younger Adults, 12 point was significantly less preferred than 14 point and 16 point, with no different between 14 point and 16 point. For older Adults, participants least preferred 12 point but preferred 16 point the most. In addition, for both Younger and Older Adults, 12 point was rated lower than midpoint while 14 point and 16 point was rated above the mid point.

Moreover, on overall preference, there were small sized interaction effects between Font Type, Font Size, Age Group, and Nationality. In order to provide information for recommendation on Font Type and Font Size for both the UK and Thai and for both the Younger and Older Adults, the results in this interactions were presented for each of the four groups of participants: Younger UK, Older UK, Younger Thai and Older Thai. For UK Younger Adults, they significantly preferred Serif 14 point and 16 point over Serif 12 point whereas they significantly preferred Sans Serif 14 point over than Sans Serif 12 point but no other significant differences. For UK Older Adults, for both Serif and Sans Serif font, they significantly preferred 14 point and 16 point over 12 point. Serif 12 point was rated below the midpoint of the rating scale, Serif 14 point and Sans

Serif 12 point was around the midpoint, while Serif 16 point, Sans Serif 14 and 16 point were rated above the midpoint. These results supported to use 14 point size and larger for UK Younger Adults and use 16 point for UK Older Adults.

For Thai Younger Adults, Sans Serif 12 point was significantly less preferred than Sans Serif 14 and 16 point. Whereas no difference on overall preference for all Font Size in Serif Font Type. Again, all Font Sizes in Serif Font Type were rate above the midpoint of the rating scale while Sans Serif 12 point was below the midpoint and Sans Serif 14 and 16 point were around the midpoint of the rating scale. For Thai Older Adults, for both Serif and Sans Serif, 12 point was significantly less preferred than 14 and 16 point. Serif 12 point was rated below the midpoint but Serif 14 and 16 point were rated above the midpoint of the rating scale. Whereas, for San Serif Font Type, all Font Sizes were lower than the midpoint, at least 16 point just reach around the midpoint. The results in line with UK participants, support to use 14 and 16 point for Thai Younger Adults, and use 16 point size for Thai Older Adults.

The current results supported the previous research conducted by Chadwick-Dias, McNulty and Tullis (2003). They also found that Older Adults preferred larger Font sizes when compared with Younger Adults. Based on these results, it can be concluded that when using a skimming task, which is representative of the type of reading that is done on the web a lot of the time, 16 point is better for older adults in the UK. Although Font size had no main effect or interaction effect on performance measures, the results on participants' preference were supported by Bernard, Liao, and Mills (2001a, 2001b) who found that, on Older Adults participants' preference, a larger font size was more preferred than a small one.

Perhaps the most interesting aspect of these results is that there is no detectable difference between 14 and 16 point by younger adults. Younger adults rated 14 and 16 point almost identically on all aspects and suffered no ill effects in performance on speed that could be detected. From a universal design point of view choosing 16 point as a minimum would have no effect on younger adults and would greatly benefit the aging population in the UK.

The study has a number of small limitations which need to be taken into account. Firstly, more of the Thai Older Adult participants, although in the right age range, were in work and therefore perhaps more likely to have experience with computer technology than the UK participants. Secondly, as noted above, although I attempted to create reading texts that would be unfamiliar but interesting to participants in both countries, the UK participants were more accurate in reading the texts. This may have

been because the topic of the texts, the Olympic Games, are more familiar to UK participants, particularly as the Olympic Games were held in London in 2012. Thus the reading materials may not have been as cultural neutral as I had hoped.

Finally, in terms of generalisability, the participants for the English language website were all English native speakers living in the UK. Whether these results can legitimately be generalised to the presentation of text on websites in languages which use the Latin alphabet other than English is unclear.

Conclusions: the recommendation on font type and font size for evidence-based research web design guidelines for older adults

Table 5.14 and Table 5.15 present the six criteria which were considered when developing the recommendations on Font Type and Font Size, respectively, for web design guidelines for older adults. From these two tables, the implications for the design of web based text content for older web users in the UK and Thailand are presented in detail, below.

Table 5.14: The six criteria which considered for making recommendation for web design guidelines on Font Type for Younger and Older Adults in the UK and Thailand

Criteria\Nationality and Age Group	UK		Thai	
	Younger Adults	Older Adults	Younger Adults	Older Adults
Time spent per webpage	**Serif** > San serif			
Overall fatigue	**Serif = San serif**			
URE	**Sans Serif** > Serif		**Serif** > Sans Serif	
URE against the mid-point on the 5 point rating scale	**Serif** ↑, **San serif** ↑		San serif ↓, **Serif** ↑	
Overall preference	**Sans serif** > Serif		**Serif** > Sans serif	
Overall preference against the mid-point on the 5 point rating scale	Serif ↓, **San serif** ↑		San serif ↓, **Serif** ↑	

A > B means A significantly better than B , = means no significant difference

- means rating is not significantly different from mid-point, ↑ means rating is significantly above mid-point, ↓ means rating is significantly lower than mid-point

* the bold text means it was recommended in that criteria

Table 5.14 shows that most of criteria recommend San Serif Font Type for UK Younger and Older Adults. On the other hand, all criteria recommended Serif for Thai Younger and Older Adults.

The recommendations on Font Type for the UK older people from the current study support to use the use of Sans Serif fonts. This recommendation is same as recommendations of SPRY Foundation (Holt and Komlos-Weimer, 1999), Holt (2000), The National Institute on Aging/National Library of Medicine (NIA/NLM) (2002), ARRP (2004) and the SilverWeb guideline (Kurniawan and Zaphiris, 2005; Zaphiris, Kurniawan and Ghiawadwala, 2007). In addition, Sans Serif is also recommended for UK Younger Adults, so this is not a recommendation particularly for older web users, but for all the UK web users.

Table 5.15 shows that most of the criteria recommended 14 or 16 point for the UK and Thai Younger Adults whereas 16 point was recommended for the UK and Thai Older Adults.

This study has shown that 12 point does not provide an optimal user experience for either UK Younger and Older Adults in any way. For creation of web text content, a minimum of 14 point should be recommended for UK younger adults and 16 point for UK older adults. Indeed, the argument can be made that 16 point font has no negative impacts on younger adults and therefore could be used as a minimum font size in any future guidelines for older adults. Thus, the current study does not support any previous recommendations.

Table 5.15: The six criteria which considered for making recommendation on Font Size for Younger and Older Adults in the UK and Thailand

Criteria\Nationality and Age Group		UK		Thai	
		Younger Adults	Older Adults	Younger Adults	Older Adults
Time spent per webpage		no significant difference			
Overall fatigue		12>**16**, 12 = 14, **14=16**			
URE		**14,16** > 12	**14,16** > 12	**14,16** > 12	**14,16** > 12
URE against the mid-point on the 5 point rating scale		12↓, **14**↑, **16**↑	12↓, 14 -, **16**↑	12-, **14**↑, **16**↑	12↓,**14**↑,**16**↑
Overall preference	Serif	**14, 16** >12	**14, 16** >12	**12 = 14 = 16**	**14, 16** >12
	Sans Serif	**14** > 12, 12=16, 14=16	**14, 16** >12	**14, 16** >12	**14, 16** >12
Overall preference against the mid-point on the 5 point rating scale	Serif	12↓, **14**↑, **16**↑	12 ↓,14 -, **16**↑	**12**↑, **14** ↑, **16** ↑	12↓,**14**↑,**16**↑
	Sans Serif	12 -, **14**↑, **16** ↑	12 -, **14** ↑, **16** ↑	12 ↓,**14** -, **16** -	12↓, 14 ↓, **16-**

A > B means A significantly better than B, = means no significant difference

- means rating is not significantly different from mid-point, ↑ means rating is significantly above mid-point, ↓ means rating is significantly lower than mid-point

* the bold text means it was recommended in that criteria

Moreover, the current study suggests the first recommendations for Font Type for Thai Younger and Older Adults which are totally different from the recommendation for UK Younger and Older adults. Thai conservative Font Type, which corresponds most closely to a serif font in the Latin alphabet, is recommended. For Font size, 16 point is recommended for Thai older web users. However, text with 14 point is still recommended for Thai younger web users, see Table 5.16.

Table 5.16: The recommendation on font type and font size for evidence-based research web design guidelines

Age group and Nationality	Recommendations on font type	Recommendations on font size
UK younger adults	san serif	14 point or larger
UK older adults	san serif	16 point
Thai younger adults	Thai conservative (serif)	14 point or larger
Thai older adults	Thai conservative (serif)	16 point

Perhaps the most important conclusion that can be drawn from this study is that that a single set of guidelines on font type cannot apply across multiple languages, and in particular those with a different alphabet. Either due to differences in the fonts or due to differences in the participants in the different nations, the results are very different for Latin and non-Latin fonts.

Chapter 6

Effects of text colour and background colour on skim reading web pages by younger and older people in the UK and Thailand

---

6.1 Introduction

This chapter presents the fourth study in my programme of research, which investigated the effect of text colour and background colour on skim reading webpages by younger and older adults both in Thailand and the UK.

Text colour and background colour were chosen to investigate as independent variables in this study because most of the web design guidelines provide recommendations about text colour and background, as dicussed in the literature review in Chapter 2 (see section 2.7 in particular). However, the recommendations they make about this aspect are varied and very little empirical evidence is provided to support them. Thus it is unclear which particular text colour and background colour combinations would be better for older adults.

There are eight sets of guidelines which make recommendations text colour and background colour: SPRY (1999), Holt (2000), Zhao (2001), Agelight (2001), NIA/NLM (2002), AARP (2004), SilverWeb (2005, 2007), and Webcredible (2006). Six of these sets of guidelines recommend using dark text on a light background (SPRY, 1999; Holt, 2000; Zhao, 2001; Agelight, 2001; NIA/NLM, 2002; Webcredible, 2006). Two sets of guidelines also mention that "reverse contrast" can be used, a light text on a dark background (Holt, 2000; NIA/NLM, 2002). Two set of guidelines mention only that strong contrast between text and background should be used (ARRP, 2004; SilverWeb,

2005, 2007). Two sets of guidelines also mention that off-white rather than pure white is a better background (SilverWeb, 2005, 2007; Webcredible, 2006).

Currently, there is no specific recommendation on text colour and background colour for the Thai language web site or any web design guidelines for Thai younger or older people.

This current study used a range of combinations of text colour and background colour to present content on a web site. The participants were both older and younger people in both the UK and Thailand. The UK participants participated using an website in English while the Thai participants participated using the same website, but translated into Thai. This range of participants allowed me to investigate which of the tested combinations is most appropriate for older web users in both Thailand and the UK.

The following research questions were addressed by this study:

1. Does text colour and background colour have an effect on reading performance and preferences of younger and older people when reading web pages?

2. Does age group have an effect on reading performance and preferences when reading web pages?

3. Does nationality have an effect on reading performance and preferences when reading web pages?

4. Do attitudes toward the web have an effect on reading performance when reading web pages?

6.2 Method

6.2.1 Design

A four way mixed design was used in this experiment. Age Group and Nationality were the between participant independent variables, and Text Colour and Background Colour Combinations (Text/Background Colour) and Task (each participant did two tasks with each text/background colour combination) were the within participant independent variables.

The independent variables had the following levels:

- Age Group - participants were either Older Adults (55 years and over for participants in Thailand and 65 years and over for participants in the UK, see Section 2.3 for calculation of appropriate minimum age for older adults in Thailand and the UK) or Younger Adults (18 to 39 years in both Thailand and the UK)

- Participants nationality, language and writing system (Nationality for short) - participants were either British people in the UK who were native speakers of English or Thai people in Thailand who were native speakers of Thai

- Text Colour and Background Colour Combinations (Text/Background Colour) - either Black text on White Background (Black/White), White text on Black Background (White/Black), or Sepia Text on Off-White Background (Sepia/Off-white)

- Task - task 1 and task 2

Black text on white background was selected as most of computer systems present text in this combination of colours and it is the same as used very frequently in print media. White text on black background was selected as it is recommended by a number of sets of web design guidelines for older people (Holt, 2000; NIA/NLM, 2002). Sepia text on off-white background was selected as it is a colour combination which is provided to use for reading in some computer systems such as e-readers (iBook, readMe, eBookMobile, NeoBook, and eReader). Figures 6.1 - 6.3 show examples of combinations of text and background colours which were used in the current study.
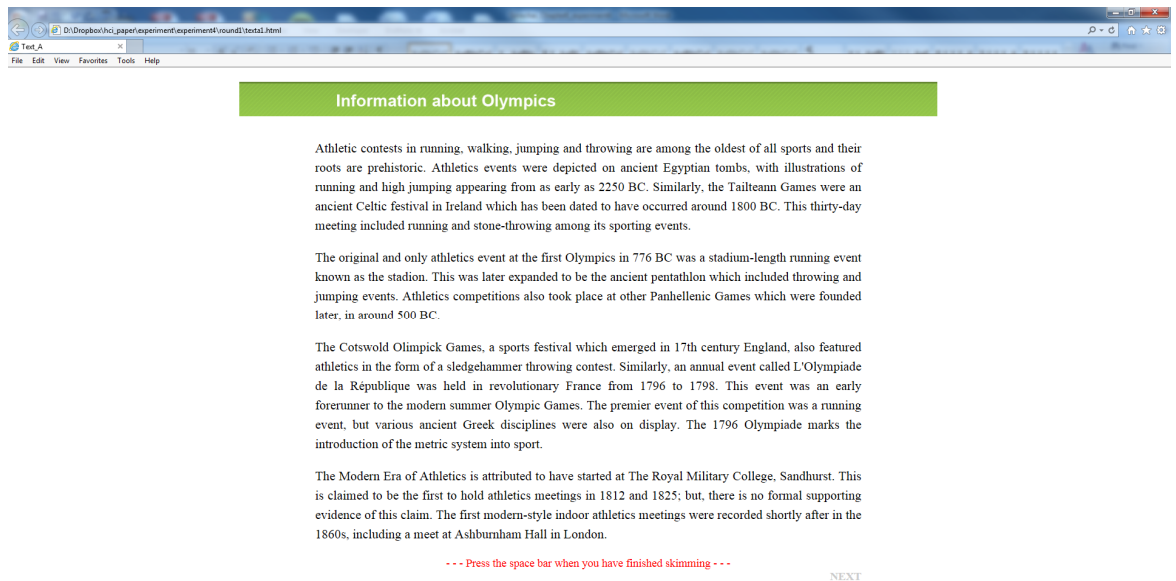
Figure 6.1: Examples of Black Text on White Background Combination Condition
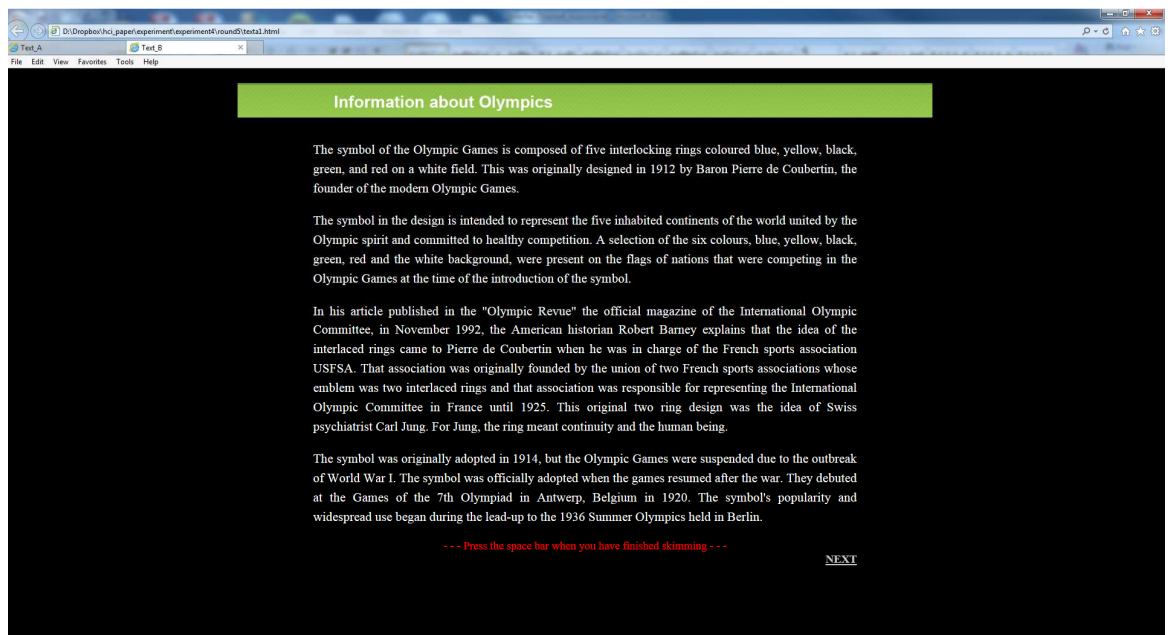
Figure 6.2: Examples of White Text on Black Background Combination Condition
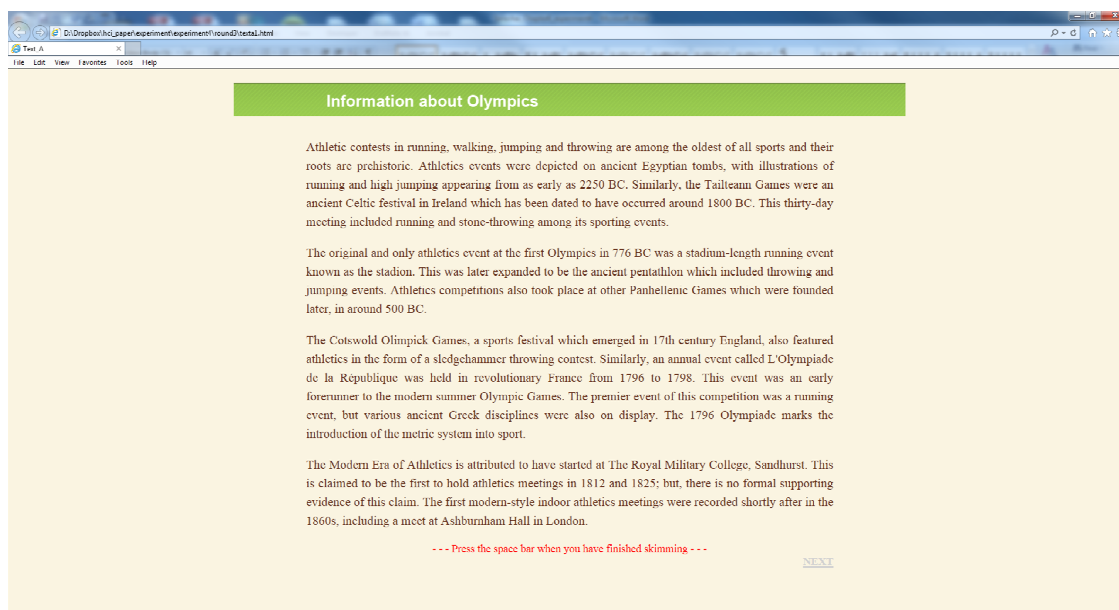
Figure 6.3: Examples of Sepia Text on Off-white Background Combination Condition

Each participant undertook six tasks on web pages, two tasks with each of the three combinations of Text/Background Colour. Tasks were to skim read the text and answer four multiple choice questions, the same task as used successfully in the previous experiment. The order of presentation of the six tasks was counterbalanced between participants to compensate for practice and fatigue effects.

Two dependent variables related to participants' performance were measured:

- Time spent per web page – this was a measure of the reading speed of participants on the web page.

- Number of correct answers – this was measure of the efficiency. The number of correct answers were calculated and converted to the percentage of correct answer before analyses were undertaken

Four dependent variables related to participants' preferences were measured:

- Participants' attitude towards the web were measured using the Attitudes toward the Web Scale (ATWS) (Burn, 2003) which has three factors: Confidence, Performance, and Fashion, (see Chapter 2, Section 2.4 for more details).

- Participants' rating of visual and physical fatigue (adapted from Tyrrell and Leibowitz, 1990)

- Participants' ratings of each Text/Background Colour combination condition on three dimensions: Ease of Reading; Pleasantness of Reading; and Speed of reading

- Participants' ratings of their overall preference of each of combination of Text/Background Colour

The dependent variables above are those usually measured in research about reading text (see Dillon, 1992, 2004). All the preference dependent variables were measured on five point Likert items, except the participants' rating on the visual and physical fatigue which was measured on seven point Likert items. The seven point Likert item was adopted from Dillion, Kleinman, Ok Choi, and Bias (2006).

6.2.2 Participants

63 people participated in this current study. This included 27 participants in the UK and 36 participants in Thailand. All participants had not participanted in the second or third studies this programme of research.

The UK participants comprised 18 Younger and 9 Older Adults. The UK Younger Adults comprised 12 males and 6 females, aged between 18 and 36 years (Mean=23.72 years, SD=5.52). There were 11 undergraduate students, one Masters degree student, five Ph.D. students, and one person who was employed. The UK Older Adults comprised 4 males and 5 females, aged between 66 and 79 years (Mean=73.56 years, SD=4.19). All the UK Older participants were retired.

The Thai participants comprised 18 Younger and 18 Older Adults. The Thai Younger Adults comprised 7 males and 11 females, aged between 19 and 29 years (Mean=23.72 years, SD=3.06). There were 11 undergraduate students, one Masters student, three Ph.D. students and three employed persons. The Thai Older Adults comprised 3 males and 15 females, aged between 59 and 70 years (Mean=61.33 years, SD=3.07). 12 were employed and the other six were retired.

The Younger Adults participants were offered a gift voucher valued at £10. The Older Adults participants were offered a gift voucher valued at £15, as the sessions for the Older Adults took considerably longer than those for Younger Adults.

The UK Younger Adult participants were recruited through announcements on the University of York Graduate Student Association web site, and print advertisements posted at university bus stops.

The UK Older Adult participants were recruited through the York Older People's Assembly and the panel of older people who had participated in previous studies for the Human Computer Interaction Research Group in the Department of Computer Science at the University of York.

The Thai Younger Adult participants were recruited from students, staffs, and alumni of Suranaree University of Technology. The Older participants were recruited from lecturers from Nakhon Ratchasima Vocational College and Rajamangala University of Technology Isan.

In both the UK and Thailand, a snowball recruiting strategy was also used with the Older Adult participants, once someone had taken part in the study, they were asked to ask their friends if they would like to participate in the study.

6.2.3 Equipment and materials

Equipment and materials in this study were the same as used in the third study presented in Chapter 5. In line with the results of the first and the third studies, the content was presented as 16 point with 1.5 line spacing and left - right justification. The text in English was Time New Roman font type while the text in Thai was Conservative font type, both were serif.

The text and background colours in this study were black (#000000), white (#FFFFFF), sepia (#5E2612), and off-white (#F5EFDC). The brightness difference, the colour difference, and the contrast ratio between Black / White were 255, 765, and 21, respectively. The brightness difference, the colour difference, and the contrast ratio between White / Black were 255, 765 and 21, respectively. The brightness difference, the colour difference, and the contrast ratio between Sepia / Off-white were 186.16, 554, and 10.37, respectively.

6.3 **Procedure**

The study was conducted at a number of locations, all quiet rooms at the institutions where the participants studied or where they came to take part in the study.

The procedure in this study was the same as the experiment in Chapter 5.

Each session took approximately 20 minutes to complete for Younger Adult participants, and approximate 40 minutes to complete for Older Adult participants.

6.4 Data preparation

Firstly, I did the histograms from the Data on Time spent per webpage in each combinations of combination of Text/Background Colour. The histogram was not a normal distribution, it was necessary to normalise before doing the analysis by using the method outlined in Section 3.4.1. In this process, 17 data points out of a total of 396 data points (4.29%) were adjusted.

6.5 Results

Analysis of Variance (ANOVA) was used to investigate the effects of the independent variables of combinations of Text/Background Colour, Age Group and Nationality on both performance and preference measures and the appropriate post-hoc analyses were conducted when there were any significant effects from the overall ANOVA analysis. In addition, for participants' preference ratings, t-tests were used to investigate whether preference ratings were significantly above or below the mid-point of the rating scale.

6.5.1 Time spent per web page

A four way Analysis of Variance (ANOVA) of Text/Background Colour by Age Group by Nationality by Task on the time spent per webpage found that Age Group and Nationality had significant effects (Age Group: $F(1,62)=6.02$, $p<.05$, $\eta_p^2=0.09$; Nationality: $F(1,62)=10.44$, $p<.01$, $\eta_p^2=.14$). However, there were no significant effects of Text/Background Colour combinations ($F(2,124)=2.12$, n.s.) or Tasks ($F(1,62)=2.14$, n.s.). In addition, there was a significant interaction between Age Group and Nationality $F(1,62)=4.05$, $p<.05$, $\eta_p^2=.06$). There were no other significant interactions effects.

Older Adults spent significantly longer time per web page than Younger Adults (Mean Younger Adults=62.87 sec SD=26.62; Mean Older Adults=76.93 sec SD=24.45).

The UK participants spent significantly shorter time per web page than the Thai participants (Mean UK participants=58.30 SD=22.49; Mean Thai participants=78.39 SD=26.19).

A Scheffé post hoc analysis was used to test the specific differences in the interaction between Age Group and Nationality. Table 6.1 shows the pattern of observed t-values for the interaction. The UK Younger Adults (Mean=48.55 SD=13.62) spent significantly less time per page than the UK Older Adults (Mean=72.93 SD=25.33) (Observed t=-3.75, p <.05, critical t at 95% confidence level: 3.46, at 99%: 4.60), Thai Younger Adults (Mean=77.18. SD=28.72) (Observed t=-3.96, p <.05), and Thai Older Adults (Mean= 79.59 sec. SD=23.59) (Observed t=-5.27, p<.01). There were no other significant differences. This interaction is shown in Figure 6.4.

Table 6.1: Observed t-values between all pairs of Age Group and Nationality combination for Time spent per web page

		UK		Thai	
		Younger	**Older**	**Younger**	**Older**
UK	Younger	-	-3.75*	-3.96*	-5.27**
	Older		-	-0.44	-0.81
Thai	Younger	-		-	-0.29
	Older	-			-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.46, at 99%: 4.59
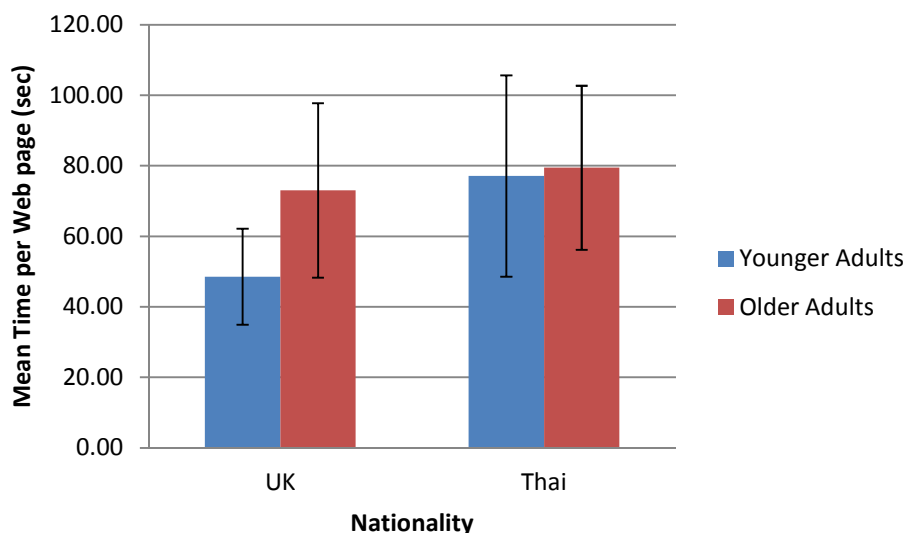
Figure 6.4: Mean Time per Web Page (seconds) for Younger and Older Adults for the UK and Thai participants

6.5.2 Percentage of correct answers

For Percentage of correct answers, an four way Analysis of Variance (ANOVA) of Text/Backgournd Colour combinations by Age Group by Nationality by Task found that Text/Background Colour combinations and Task had no significant effect on the Percentage of correct answers (Text/Background: $F(2,124)=1.66$, n.s.; Task: $F(1,62)=3.34$, n.s.). However, there were significant effects of Age Group ($F(1,62)=7.34$, $p<.01$, $\eta_p^2=.11$) and Nationality ($F(1,62)=6.60$, $p<.05$, $\eta_p^2=.10$). There were no significant interactions between any of the independent variables.

The Older Adults were significantly more accurate at answering the questions than Younger Adults (Mean Older Adults=49.17 SD=26.54; Mean Younger Adults=41.09 SD=25.94).

The UK participants were significantly more accurate at answering the questions than Thai Participants (Mean UK participants=48.89 SD=26.14; Mean Thai participants=41.32 SD=26.36).

To investigate whether there was a speed-accuracy trade-off in the way the participants undertook the tasks, the Time Spent per Web Page and the Percentage of Correct Answers were correlated. This correlation for Younger Adults was not significant (Younger Adults: $r(30) = 0.03$, n.s.) but it was for Older Adults (Older Adults:

r(30)=.46, p<.05). For older people, the participants who took longer times per web page had higher percentages of correct answers.

6.5.3 Visual and Physical Fatigue

Correlation analyses between the participants' rating on these five aspects (question 1-5) and overall fatigue (question 6) were assessed to investigate whether the participants were measuring the different things.

A strong pattern of correlations, (shown in see Appendix 17), suggests that rating on visual and physical fatigue in 5 question were not different from rating on overall fatigue in question 6. Then the overall fatigue in question 6 represented the 5 aspects of visual and physical fatigue. The overall fatigue for each participant for each of the three Text/Background Colour levels was used for the further analysis.

A three way Analysis of Variance (ANOVA) of Text/Background Colour by Age Group by Nationality on the overall fatigue rating found that Age Group and Nationality had no significant effects on overall fatigue (Age Group: $F(1,62)=1.50$, n.s.; Nationality: $F(1,62)=.02$, n.s.). However, there was a significant effect of Text/Background Colour combinations ($F(2,124)=9.12$, $p<.01$, $\eta_p^2=.13$). There was a significant interaction between Text/Background Colour combinations and Age Group ($F(2,124)=3.61$, $p<.05$, $\eta_p^2=.06$). There were no other significant interactions.

A Scheffé post hoc analysis was used to investigate the specific differences in the interaction between Text/Background Colour combinations which is illustrated in Figure 6.5. Overall fatigue for White/Black ((Mean=3.06 SD=1.58) was significantly higher than overall fatigue for Black/White (Mean=2.42 SD=1.55) (Observed t=-3.58, p<.01, critical t at 95% confidence level: 2.48, at 99%: 3.09), or for Sepia/Off-white (Mean=2.35 SD=1.46) (Observed t=3.74, p<.01). There were no other significant differences.
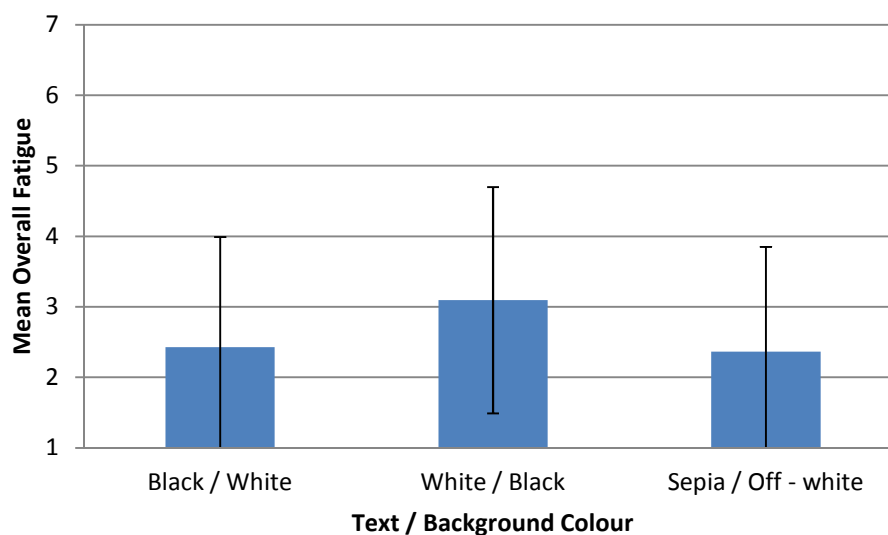
Figure 6.5: Mean overall fatigue for different Text/Background Colour combinations

One sample t-tests against the neutral mid-point rating of 4 showed that the mean overall fatigue ratings for Black/White, White/Black, and Sepia/Off-white were all significantly lower than neutral (Black/White: $t(62)=-8.26$, $p<.001$; White/Black: $t(62)=-4.84$, $p<.001$; Sepia/Off-white: $t(62)=-9.18$, $p<.001$).

A Scheffé post hoc analysis was used to investigate the specific differences in the interaction between Text/Background Colour combinations and Age Group. Table 6.2 shows the pattern of the observed t-values for the interaction. No observed t-value between any combination was greater than the critical t-Scheffe value. So the post-hoc analysis failed to reveal any significant differences in the interaction between Text/Background Colour combinations and Age Group, suggesting it was a marginal interaction. This interaction is shown in Figure 6.6.

Table 6.2: Observed t-values between all pairs of Text/Background Colour combinations and Age Group combination for Overall fatigue

		Younger Adults			Older Adults		
		Black/ White	White / Black	Sepia	Black/ White	White / Black	Sepia
Younger Adults	Black/ White	-	-2.84	1.36	1.24	0.10	0.69
	White / Black		-	3.84	3.32	2.23	2.71
	Sepia			-	0.48	-0.78	-0.09
Older Adults	Black/ White				-	-2.28	-1.06
	White / Black					-	1.13
	Sepia						-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.92
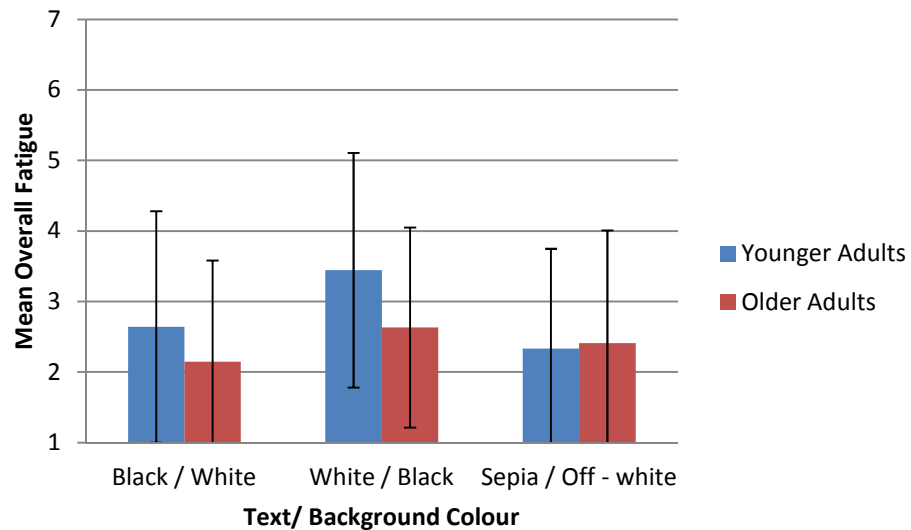
Figure 6.6: Mean overall fatigue for different Text/Background Colour combinations for Younger and Older Adults

6.5.4 Preference measures

The preference measures were participants' ratings of Ease, Pleasantness, and Speed of Reading for each Text/Background Colour combination and their overall rating of preference of each combination.

Correlations between the participants' rating on the three specific preference measures were calculated to investigate whether the ratings were measuring different dimensions of the participants' experience.

A strong pattern of correlations between all three measures for the Text/Background Colour combinations variables (see Appendix 18), meant that participants had only one underlying experience dimension on which to rate the reading tasks.

Therefore a combined User Reading Experience (URE) scores was calculated for each participant for each of the three Text/Background Colour combinations levels. This URE scores was the mean of the three ratings for each condition. An added benefit of using this combined score is that scores made up from a number of individual measures are more robust than individual items from participants (Kline, 2000).

6.5.4.1 Analysis of User Reading Experience scores (UREs) for Text/Background Colour combinations, Age Group, and Nationality

A three way ANOVA of Text/Background Colour by Age Group by Nationality on URE Scores found that Text/Background Colour combinations had a significant effect on the URE scores (Text/Background: $F(2,124)=22.78$, $p<.001$, $\eta_p^2=.27$). However there was no significant effect for Age Group, ($F(1,62)=0.73$, n.s.), or Nationality (Nationality: $F(1,62)=2.36$, n.s.). There was a significant interaction between Text/Background Colour and Age Group ($F(2,124)=6.04$, $p<.01$, $\eta_p^2=.09$). There were no other interaction effects.

A Scheffé post hoc analysis was used to test the specific differences between Text/Background Colour. The analysis found that the mean URE scores for White/Black (Mean=2.83 SD=1.04) were significantly lower than both the mean URE scores for Black/White (Mean=3.95 SD=0.70) (Observed t=7.00, $p<.01$, critical t at 95%: 2.48, at 99%: 3.09), and for Sepia/Off-white (Mean=3.39 SD=0.85) (Observed t =-3.16, $p<.01$). Mean URE scores for Sepia/Off-white (Mean=3.39 SD=0.85) were significantly lower than the mean URE scores for Black/White (Mean=3.95 SD=0.70) (Observed t=3.62, p<.01). This interaction is shown in Figure 6.7.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean rating for Black / White and Sepia/Off-white were significantly higher than neutral (Black / White: $t(65)=11.10$, $p<.001$; Sepia/Off-white: $t(65)=3.70$, $p<.001$), but the mean ratings for White/Black were not different from mid-point (White/Black : $t(65)=-1.30$, n.s.).
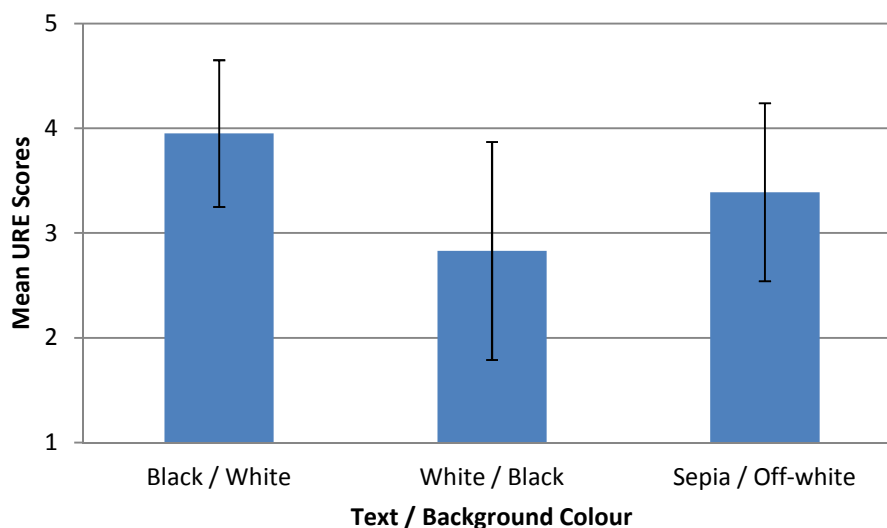
Figure 6.7: Mean URE scores for the different Text/Background Colour combinations

A Scheffé post hoc analysis was used to test the specific differences in the interaction between Text/Background Colour combinations and Age Group. Table 6.3 shows the pattern of observed t-values for this analysis.

For Younger Adults, White/Black (Mean=2.77 SD=1.15) scored significantly lower than Black/White (Mean=3.81 SD=0.61) (Observed t=4.86, p<.01, critical t at 95% confidence level: 3.92, at 99%: 4.89). White/Black (Mean=2.77 SD=1.15) scored marginally significantly lower than Sepia/Off-white (Mean=3.72 SD=0.72) (Observed t=-3.82, n.s.). But there was no difference in URE scores between Black/White and Sepia/Off-white (Mean Black/White=3.81 SD=0.61; Mean Sepia/Off-white=3.72 SD=0.72) (Observed t-values=0.52, n.s.).

For Older Adults, Black/White (Mean=4.12 SD=0.77) scored significantly higher than both White/Black (Mean=2.91 SD=0.92) (Observed t=5.00, p<.01) and Sepia/Off-white (Mean=2.99 SD=0.83) (Observed t=4.91, p<.01). But there was no difference in URE

scores between White/Black and Sepia/Off-white. This interaction is shown in Figure 6.8.

In addition, URE Scores for Black/White for Younger Adults (Mean=3.81 SD=0.61) were significantly higher than both URE scores for White/Black for Older Adults (Mean=2.91 SD=0.92) (Observed t=4.77, p<.05) and URE scores for Sepia/Off-white for Older adults (Mean=2.99 SD=0.83) (Observed t=4.63, p<.05). URE scores for Sepia/Off-white for Younger Adults (Mean=3.72 SD=0.72) were significantly higher than for White/Black for Older Adults (Mean=2.91 SD=0.92) (Observed t =4.02, p<.05).

One sample t-tests against the neutral mid-point rating of 3 showed that the mean URE scores for Black/White for both the Younger and Older Adults, and Sepia/Off-White for Younger Adults were significantly higher than neutral (Black/White Younger Adults: t(35)=7.96, p<.001; Black/White Older Adults: t(29)=8.03, p<.001; Sepia/Off-white Younger Adults: t(35)=5.98, p<.001). While the other combinations were not significantly different from neutral (White/Black Younger Adults: t(35)=-1.21, n.s.; White/Black Older Adults: t(29)=-0.54, n.s.; Sepia/Off-white Older Adults: t(29)=-0.08, n.s.)

Table 6.3: Observed t-values between all pairs of Text/Background Colour combinations and Age Group combination for User Reading Experience scores

		Younger Adults			Older Adults		
		Black/ White	White / Black	Sepia	Black/ White	White / Black	Sepia
Younger Adults	Black/ White	-	4.86**	0.52	-1.81	4.77*	4.63*
	White / Black		-	-3.82	-5.51**	-0.55	-0.86
	Sepia			-	-2.18	4.02*	3.83
Older Adults	Black/ White				-	5.00**	4.91**
	White / Black					-	-0.36
	Sepia						-

* p<.05, ** p <.01   critical t at 95% confidence level: 3.92, at 99%: 4.89
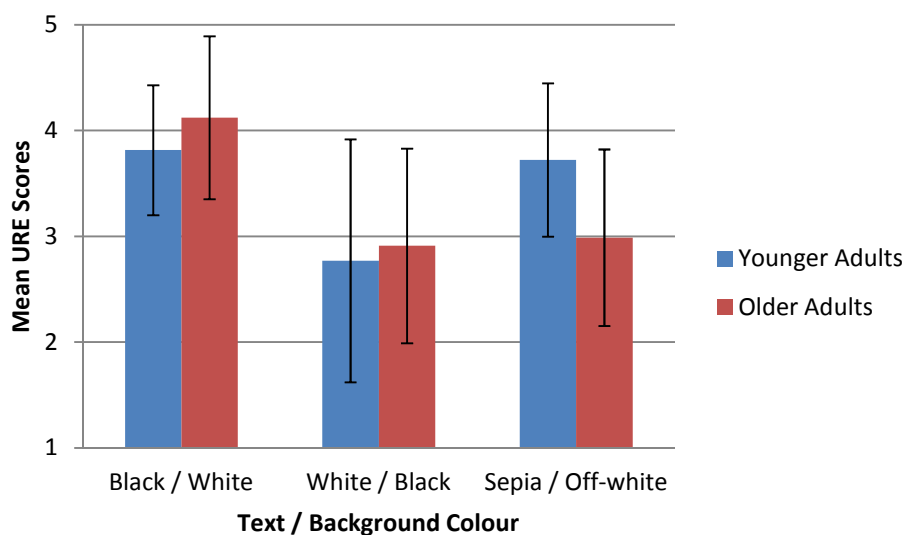
Figure 6.8: Mean URE scores for different Text/Background Colour combinations for Younger and Older Adults

6.5.4.2 Participants' overall preference ratings of combinations of Text/Background Colour

A three way ANOVA of the Text/Background Colour by Age Group by Nationality on the participants' overall preference ratings found that Text/Background Colour combinations had a significant main effect ($F(2,124)=24.70$, $p<.001$, $\eta_p^2=.29$) while Age Group and Nationality did not have significant effects (Age Group: $F(1,62)=1.57$, n.s.; Nationality: $F(1,62)=1.02$, n.s.). There was a significant interaction between Text/Background Colour combinations and Age Group ($F(2,118)=14.19$, $p<.001$, $\eta_p^2=.18$). There were no other significant interactions.

A Scheffé post hoc analysis found that overall preference ratings for Black/White (Mean=4.03, SD=0.96) were significantly higher than both White/Black (Mean=2.39 SD=1.36) (Observed t=7.48, $p<.01$, critical t at 95%: 2.48, at 99%: 3.09) and Sepia/Off-White (Mean=3.55, SD=1.34) (Observed t=-4.34, $p<.01$). Overall preference ratings for White/Black was not significantly different from Sepia/Off-White. This interaction is shown in Figure 6.9.

One sample t-tests against the neutral mid-point rating of 3 showed that the mean rating for White/Black was significantly lower than neutral ($t(65)=-3.63$, $p<.01$), but the mean ratings for Black/White and Sepia/Off-white were both significantly higher than neutral (Black/White: $t(65)=8.72$, $p<.001$; Sepia/Off-white: $t(65)=3.31$, $p<.01$).
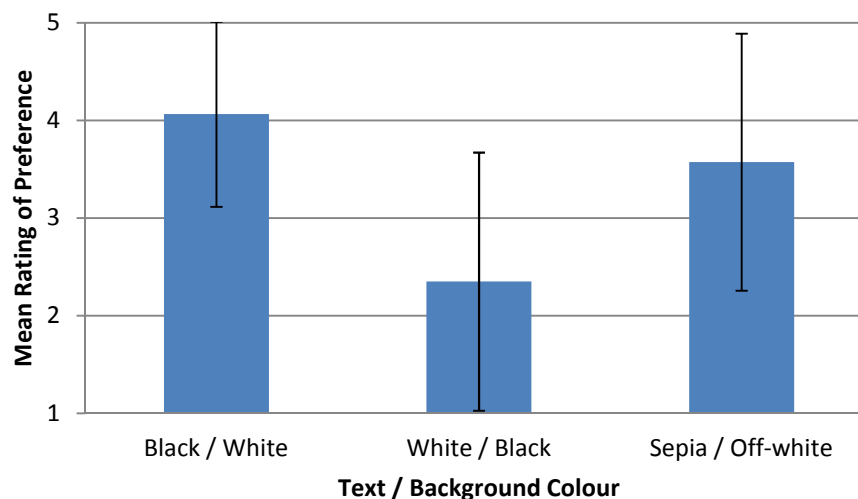
Figure 6.9: Mean Rating of Overall Preference for Text/Background Colour

A Scheffé post hoc analysis was used to test the specific differences in the two way interaction between Text/Background Colour combinations and Age Group. Table 6.4 shows the observed t-values for each combination of Text/Background Colour combinations and Age Group. For Younger Adults, White/Black (Mean=2.17, SD=1.36) was rated significantly lower than both Black/White (Mean=3.81, SD=0.82) (Observed t =6.12, p<.01, critical t at 95%: 3.92, at 99%: 4.89) and Sepia/Off-white (Mean=4.19, SD=1.06) (Observed t=-6.11, p<.01) while Black/White and Sepia/Off-white were not significantly different from each other.

For Older Adults, Black/White (Mean=4.30, SD=1.05) was rated significantly higher than both White/Black (Mean=2.67, SD=1.32) (Observed t=4.49, p<.05) and Sepia/Off-white (Mean=2.76, SD=1.22) (Observed t=4.58, p<.05) while White/Black and Sepia/Off-white were not significantly different from each other. This interaction is shown in Figure 6.10.

In addition, ratings of overall preference for Black/White for Younger Adults (Mean=3.81, SD=0.82) were rated significantly higher than both White/Black for Older Adults (Mean=2.67, SD=1.32) (Observed t=4.28, p<.05), and Sepia/Off-white for Older Adults (Mean=2.76, SD=1.22) (Observed t=4.11, p<.05). Ratings of overall preference for White/Black for Younger Adults (Mean=2.17, SD=1.36) were rated significantly lower than Black/White for Older Adults (Mean=4.30, SD=1.05) (Observed t=-6.99, p<.01). Ratings of overall preference for Sepia/Off-White for Younger Adults (Mean=4.19, SD=1.06) were significantly higher than both White/Black for Older Adults (Mean=2.67, SD=1.32) (Observed t=5.20, p<.01), and Sepia/Off-white for Older Adults (Mean=2.76, SD=1.22) (Observed t=5.07, p<.01).

One sample t-tests against the neutral mid-point rating of 3 showed that the mean rating for the White/Black for Younger Adults was significantly lower than neutral (White/Black Younger Adults: t(35)=-3.67, p<.01). While Black/White and Sepia/Off-white for Younger Adults were significantly higher than neutral (Black/white Younger Adults: t(35) =-5.88, p<.001; Sepia/Off-white Younger Adults: t(35)=6.73, p<.001). For Older Adults, Black/White was significantly higher than neutral (Black/White Older Adults: t(29)=6.75, p<.001). While White/Black and Sepia/Off-white were not significantly different from neutral, (White/Black: t(29)=-1.38, n.s.; Sepia/Off-white: t(29)=-1.05, n.s.).

Table 6.4: Observed t-values between all pairs of Text/Background Colour and Age Group combination for Participants' overall preference

		Younger Adults			Older Adults		
		Black/ White	White / Black	Sepia	Black/ White	White / Black	Sepia
Younger Adults	Black/ White	-	6.12**	-1.39	-2.14	4.28*	4.11*
	White / Black		-	-6.11**	-6.99**	-1.51	-1.86
	Sepia			-	-0.40	5.20**	5.07**
Older Adults	Black/ White				-	4.49*	4.58*
	White / Black					-	-0.29
	Sepia						-

* p<.05, ** p <.01 critical t at 95% confidence level: 3.92, at 99%: 4.89
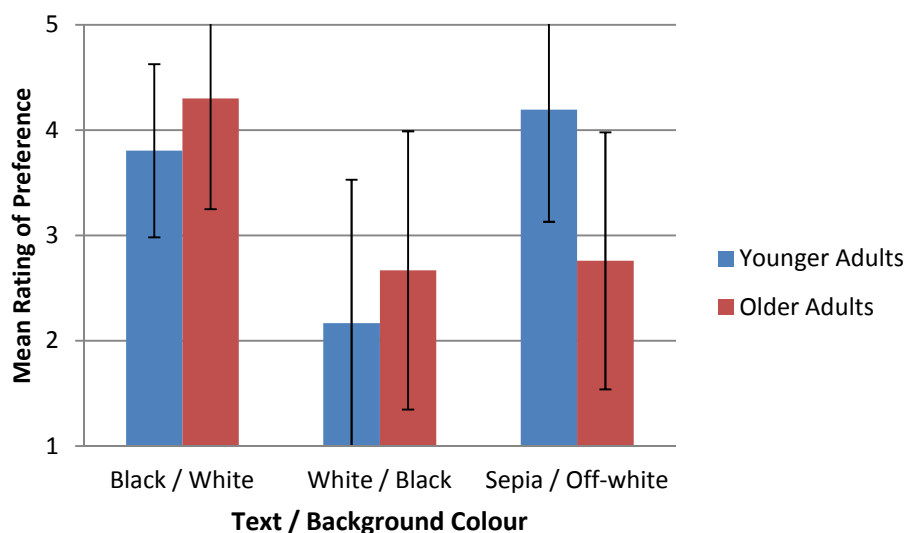
Figure 6.10: Mean Rating of Overall Preference for different Text/Background Colour combinations for Younger and Older Adults

6.5.5 Predicting reading performance from Attitudes to the Web Scale (ATWS)

A linear regression was computed to predict participants' reading performance as measured by time spent per web page from the three Factors of the ATWS, Age Group, and Nationality. The linear regression followed the formula discussed in section 3.4.5.

The linear regression produced an overall significant predication (R^2=0.31 Adjusted R^2=0.25, $F_{(5,65)}$=5.29, p<.001). The Performance Factor of the ATWS was the only significant individual predictor variable (Performance Factor: t=-2.05, p<.05), as shown in Table 6.5.

Table 6.5 The B-values, t-values, and significance levels for the linear regression predicting time per web page from the ATWS factors, Age Group, and Nationality

	B	**t**	**Sig.**
Confidence Factor	-4.42	-0.67	.51
Performance Factor	-14.62	-2.05	.045
Fashion Factor	6.35	1.58	.12
Age Group	10.72	1.63	.11
Nationality	13.81	1.89	.06

Figure 6.11 shows a scatterplot of the correlation between Performance Factor and Reading time per web page. The higher rating on Performance Factor, the less Reading time per web page.
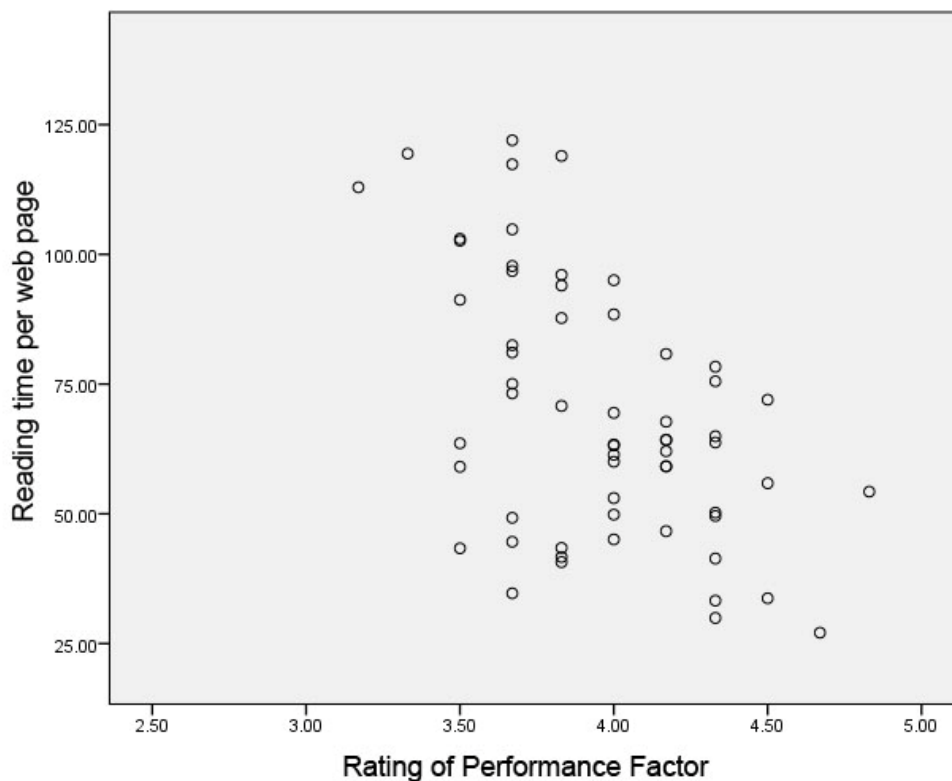
Figure 6.11: A scatterplot of the correlation between Performance Factor and Reading time per web page.

6.6 Discussion

This study investigated the effect of Text/Background Colour on performance and preference measures of skim reading webpages for Younger and Older Adults in the UK and Thailand.

On the performance measures, the first thing to consider is whether participants were changing their skim reading behaviour using a speed-accuracy trade-off. To investigate this, I asked participants four multiple choice questions about each page they skim read. There were no significant differences in the accuracy of their answers

due to Text/Background Colour or Task, although there was a small effect size[1] significant effect of Age Group and Nationality on both accuracy and skim reading time. UK participants spent shorter time but had more accurate than Thai Participants. Younger Adults read faster than Older Adults but were less accurate than Older Adults. The reason that UK participants read faster but were more accurate may have been because UK participants would have been more familiar with material about the Olympic Games than the Thai participants and found it easier to understand and assimilate new material. While the reason that Older Adults were more accurate than Younger Adults, may have been a speed-accuracy trade-off, as there was correlation between the time spent per page and the percentage of correct answers for Older Adults. Older participants who spent longer on pages were more accurate, but there was no such correction for Younger Adults. In addition, there was small significant interaction between Age Group and Nationality on Time spent per page. UK Younger Adults read faster than other groups of participants.

On the preference measures, Text/Background Colour had a small but significant effect on Overall fatigue, with White/Black having significantly higher Overall fatigue scores than both Black/White and Sepia/Off-white. On URE scores, Text/Background Colour had a small significant effect with Black/White the most preferred while White/Black was the least preferred. URE scores Black/White and Sepia/off-white were significantly above the midpoint of the rating scale and thus acceptable to participants, whereas White/Black was neutral. There was also a significant interaction between Text/Background Colour and Age Group. Younger Adults preferred Black/White over White/Black. White/Black also scored lower than midpoint of the rating scale but Black/White and Sepia/Off-white scored above the midpoint. Older Adults preferred Black/White over both White/Black and Sepia/Off-white with only Black/white scored above the midpoint, whereas White/Black and Sepia/Off-white were around the neutral. There were no different between Younger and Older Adults on their rating on each Text/Background Colour combination.

On Overall preference, the results were in line with the URE scores, Text/Background Colour a small but significant effect. There was a small significant interaction between Text/Background Colour and Age Group. Black/White was more preferred than White/Black and Sepia/Off-white. Rating on Overall preference for Black/White and

---

[1] A small effect is one that captures about 1 percent of the variance. In term of standardised difference, a small effect has value approximate 0.25. A medium effect captures about 6 percent of the variability, value is approximate 0.5. A large effect captures at least 15 percent of variability, value is approximate 0.8 (Keppel, 2004).

Sepia/off-white were significantly above the midpoint of the rating scale, but the rating for White/Black was lower than the midpoint. Younger Adults preferred Black/White and Sepia/Off-white over White/Black. Again, White/Black scored lower than midpoint of the rating scale but Black/White and Sepia/Off-white scored above the midpoint. Older Adults preferred Black/White over both White/Black and Sepia/Off-white with only Black/white scored above the midpoint, whereas White/Black and Sepia/Off-white were around the neutral. Younger participants preferred Sepia/off-white significantly more than Older participants while no different on Black/White and White/Black.

Thus, Older Adult participants preferred on Black/White but Younger Adult participants preferred Black/White and Sepia/Off-white. Older Adults participants gave some feedback after complete the study that the contrast of Sepia/Off-white was not making text clear enough for them. However, for Younger Adult participants, they said that Sepia/Off-white is good for reading for a long time. In this case, many Thai Younger Adult participants said that they are not familiar with Sepia/Off-white but the results show that they still prefer this colour combination.

From the results of the current study, White/Black is not recommended even White/Black and Black/White have the same contrast ratio but the participants felt that White/Black is too strong contrast. This makes participants more fatigued, it is shown in Overall fatigue that White/Black scored higher than other combinations of colours.

Finally, in terms of generalisability, the participants for the English language website were all English native speakers living in the UK. Whether these results can legitimately be generalised to the presentation of text on websites in languages which use the Latin alphabet other than English is unclear.

Conclusions: recommendation on text colour and background colour for evidence-based research web design guidelines for older adults in the UK and Thailand

Table 6.6 shows the six criteria which were considered for making the recommendations on Text/Background Colour combinations. From the table, the implications for design of web based text content for web readers in the UK and Thailand are presented in topic below.

Table 6.6: The six criteria which considered for making recommendations on Text colour and Background colour for Younger and Older Adults in the UK and Thailand

Criteria\Nationality and Age Group	Younger Adults		Older Adults	
	UK	Thai	UK	Thai
Time spent per webpage	no significant difference			
Overall fatigue	White/Black > **Black/White, Sepia**			
URE	**Black/White** > White/Black **Black/White = Sepia**		**Black/White** > White/Black, Sepia	
URE against the mid-point on the 5 point rating scale	White/Black -, **Black/White ↑** **Sepia ↑**		White/Black -, **Black/White ↑** Sepia -	
Overall preference	**Black/White, Sepia** > White/Black		**Black/White** > White/Black, Sepia	
Overall preference against the mid-point on the 5 point rating scale	White/Black ↓, **Black/White ↑** **Sepia ↑**		White/Black -, **Black/White ↑** Sepia -	

A > B means A significantly better than B , = means no significant difference

- means rating is not significantly different from mid-point, ↑ means rating is significantly above mid-point, ↓ means rating is significantly lower than mid-point

* the bold text means it was recommended in that criteria

Table 6.6 shows that all criteria which related to preference measures recommended both Black/White and Sepia/Off-white for both the UK and Thai Younger Adults. While, all criteria which related to preference measures recommended only Black/White for both the UK and Thai Older Adults.

Thus, the recommendations in the current research for the UK and Thai older web reader is black text on white background. This recommendation support SPRY Foundation (Holt and Komlos-Weimer, 1999), Holt (2000), Zhao (2001), Agelight (2001), and National Institute on Aging/National Library of Medicine (NIA/NLM) (2002) which recommended dark text on light or white background for older people. The current results do not support SPRY Foundation (Holt and Komlos-Weimer, 1999) and The National Institute on Aging/National Library of Medicine (NIA/NLM) (2002) which suggested light or white text on black or dark background for presenting text on webpages for older people.

Moreover, the current research suggests back text on white background and sepia text on off-white background for both the UK and Thai younger adults, see table 6.7.

Table 6.7: Recommendations on text colour and background colour for evidence-based research web design guidelines for younger and older adults in the UK and Thailand

Age group and Nationality	Recommendations on text colour and background colour
UK younger adults	Black text on white background or sepia text on off-white background
UK older adults	Black text on white background
Thai younger adults	Black text on white background or sepia text on off-white background
Thai older adults	Black text on white background

For recommendations on text colour and background colour, this study has shown that there are different recommendations for different age groups. The differences of languages and writing systems seem to have no effect. Although, sepia text on off-white background is good for younger web readers, black text on white background is good for web readers of all ages.

Chapter 7

Overall Discussion and Conclusions

7.1 Overview of the programme of research

The number of older people across the world is increasing. It is predicted that by 2050, it will be the first time in history that the proportion of the population aged 60 years and over will be larger than the proportion of people aged under 15. The situation in both the UK and Thailand is the same as in other countries across the world. This is one of the reasons that is leading to a increasing number of older web users. However, age-related physical, sensory, and cognitive capabilities of older people are potentially barriers to the use of the web by older people. In addition, some older people are less familiar than younger people are with Information and Communication Technologies, especially the web. Hence, web accessibility and usability are important issues as they will empower and support older people to be able to use and take the advantage of the web.

After reviewing the current web design guidelines for older people, I found that there are at least nine sets of web design guidelines. Most of the guidelines provide the recommendations in the English language, except one that is in Portuguese. From the review of the web design guidelines for older people there are four major issues that emerge. Firstly, what is the evidence which supports each recommendation in the web design guidelines? It was found that most web design guidelines lack evidence-based research to support their recommendations. Secondly, which recommendations in which web design guidelines are the most appropriate to apply to the design of the web for older people? This is a major issue because I have found that different web design guidelines often provide different recommendations. Thirdly, as mentioned, most of the web design guidelines are in English, it is unclear whether the recommendations in these web design guidelines for older people are appropriate only for English, only for languages using the Latin alphabet or whether web design guidelines for the Latin Alphabet can be applied for other languages, such as Thai. Lastly, the web design guidelines were suggested during the period 1999 - 2007 and the 2007 set of

guidelines contains the same guidelines which were published in 2005. So all the guidelines are at leas 10 years old. However, the web is a rapidly changing environment and the devices we use to access the web are also rapidly changing, so new evidence is needed to support guidelines.

The main goal of this research programme was to investigate a selection the recommendations from web design guidelines for older people, taking an empirical approach with older participants in Thailand and the UK. In particular, the recommendations related to text for reading web pages (line spacing, text justification, font type, font size, text colour and background colour) were investigated, using both performance and preference measures in order to create new evidence-based web design guidelines for older adults in the UK and Thailand.

In order to fulfill these goals, I have conducted four studies. The first study investigated the effect of line spacing and text justification on reading web pages by younger and older people in the UK and Thailand. As interesting issues about the task emerged in the first study (see Chapter 3 for details), the second study explored the range and appropriateness of dependent variables in reading on the web. From the second study, the skimming task was chosen as a tool to use in further research (see Chapter 4 for details). The third study investigated the effect of font type and font size on skim reading web pages by younger and older people in the UK and Thailand (see Chapter 5 for details). The last study investigated effect of text colour and background colour on skim reading web pages by younger and older people in the UK and Thailand (see Chapter 6 for details).

7.2 Implications and contributions of this thesis

7.2.1 Evidence-based research web design guidelines for design of web based text content for web readers in the UK

The first contribution of this thesis is providing a body of evidence-based research web design guidelines for design of web based text content for web readers, especially older adults, in the UK. There are a large number of web design guidelines for older people. The first set of guidelines was proposed in 1999 and the latest set of guidelines was proposed in 2007. Eight sets of web design guidelines for older people are proposed for languages written in the Latin alphabet. Most of the guidelines lack evidence-based research to support their recommendations and have been produced by reviewing previous guidelines. Some guidelines have taken the recommendations

for reading text on paper and suggested their use for the web, which may not b
appropriate.

Based on three studies in this research programme, the current evidence-based web
design guidelines recommend san serif text with 1.5 or double line spacing with either
left only or left-right justification for both the UK younger and older adults. For the
creation of web text content, a minimum of 14 point should be used for UK younger
adults while 16 point is recommended for UK older adults. Indeed, the argument can be
made that 16 point font has no negative impacts on younger adults and therefore could
be used as a minimum font size in any future guidelines for both younger and older
adults.

**Table 7.1: The evidence-based research web design guidelines for design of web
based text content for web readers in the UK**

	UK younger adults	UK older adults
Line spacing	1.5 or double line spacing	
Text justification	Left or Left-Right Justification	
Font type	San serif	
Font size	14 point or larger	16 point
Text colour and background colour	Black text on white background or sepia text on off-white background	Black text on white background

All the participants for the English language websites were native British English
speakers, so this potentially means that the results cannot be generalised across all
English speaking countries. However, there is no reason to expect that these results do
apply to all English speaking countries in the developed world where language training,
experience and exposure to the web are similar to those in the UK. Whether these
results can be generalized to other languages written in the Latin alphabet is less clear
and some research to confirm this is required.

**7.2.2 The evidence-based research web design guidelines for design of web
based text content for web readers in Thailand**

As there are no web design guidelines for both younger and older people in Thailand,
the second contribution of this thesis is providing both evidence-based research for

web design guidelines for the design of web based text content for web readers, especially older adults, in Thailand. Furthermore, it provides the first set of experiments related to text presentation on screen in the Thai language which measure both performance and preference measures.

The current evidence-based web design guidelines recommend Thai conservative text with 1.5 or double line spacing with left-right justification for both younger and older web users in Thailand. For the creation of web text content, a minimum of 14 point should be used for Thai younger web users and 16 point for Thai older web users. Indeed, the argument can be made that text with 16 point could be used in future web design guidelines for both younger and older Thai adults.

Table 7.2: The evidence-based research web design guidelines for design of web based text content for web readers in Thailand

	Thai younger adults	**Thai older adults**
Line spacing	1.5 or double line spacing	
Text justification	Left - right justification	
Font type	Thai conservative (serif)	
Font size	14 point or larger	16 point
Text colour and background colour	Black text on white background or sepia text on off-white background	Black text on white background

7.2.3 Analysis of the extent to which recommendations can be generalised across different nationalities and languages

For the third contribution of this thesis, I argue that aspects of web design guidelines might involve the same recommendations for people of same age group but with different nationalities and languages. For example, the recommendation about text colour and background colours for younger adults in the both UK and Thailand is black on white and sepia on off-white, while the recommendation for both UK and Thai older people is black on white.

It is also important to investigate which aspects of web design guidelines might be the same in recommendations for people of different age groups (but of the same nationality). Thus my research has shown that for example, the recommendations on

text justification and font type show that both the UK younger and older web users have the same requirements. Turning to Thai younger and older adults, the same recommendations on text justification and font type also appear. The most interesting recommendation on text presentation is the recommendation on line spacing. The results in this thesis show that both younger and older adults in both the UK and Thailand prefer 1.5 or double line spacing.

All these interesting issues cannot be definitely decided upon simply from the results of the one set of experiments conducted in this programme of research. Further research should be undertaken, both with the same variables used in this thesis to confirm the effects, and also with participants reading in other languages, both those using the Latin alphabet and those reading in other writing systems. Such research will show the appropriate level of generalization that can be made from experiments conducted in English.

7.2.4 Proposal of skim reading task as an appropriate task for conducting further research on reading from the web

The fourth contribution of this thesis is the proposal of skim reading as an appropriate task for carrying out research about reading text on webpages. The proposal was made based on considerations on four dimensions (see Chapter 4 for details): how time consuming a task is; how much fatigue, both visually and physically, it puts on the participant; the ability of the task to generate different reading rates and thus allow discrimination in reading performance; and finally how ecologically valid the task is in relation to web users' actual behaviour. The skim reading task satisfied all these four dimensions.

Experiment 3 (Chapter 5) and 4 (Chapter 6) used the skim reading task and the task lead to interesting results. I have noted that the skim reading task is an appropriate task for doing research about reading text on web pages. If the content in different languages are equivalent, the skimming task is also appropriate to conduct studies with people in different countries and languages.

7.2.5 Confidence and Performance Factors of the Attitudes to the Web Scale for predicting reading performance

The results from three different experiments show that the Confidence Factor and Factors of the Attitudes to the Web Scale (Burn, 2003) are useful for predicting reading time per web page. However, not surprisingly the Fashion Factor of the Attitudes to the

Web Scale does not play an significant role in predicting reading behaviour, (see Table 7.3).

Thus, the Attitudes to the Web Scale is suggested to use in the further research about reading on the web.

Table 7.3: The level of significant difference *(p* value) of the correlation between reading time per web page with Confidence Factor, Performance Factor, Fashion Factor, Age Group, and Nationality

	level of significant difference *(p* value)		
	Experiment 1 Line Spacing and Text Justification	**Experiment 3 Font Type and Font Size**	**Experiment 4 Text Colour and background Colour**
Confidence Factor	0.02	0.02	-
Performance Factor	-	0.06	0.05
Fashion Factor	-	-	-
Age Group	0.00	-	-
Nationality	-	0.00	0.06

7.3 Future work

Investigating the possibilities of interaction between text presentation variables

The experiments in the current research programme have each investigated the effects of one or two text presentation variables. For example, the first experiment investigated the effects of line spacing and text justification. The results of each experiment were then used in subsequent experiments to create optimal presentation of the text. For example, when I conducted the experiment on font type and font size, the texts were all presented using the line spacing and text justification results from the first experiment. However, there could be interactions between variables that were investigated in separate experiments in this research programme. Future research should investigate the possibilities of further interaction between text presentation variables which were not combined in the current programme of research.

Investigating more levels of each text presentation variable

Although the current research programme has made recommendations about a number of text presentation variables, it was only possible to investigate a limited number of levels for each of these variables.. For example, the study of font size in this thesis has investigated three levels: 12 point, 14 point, and 16 point. For older adults, the recommendation is that 14 point is good and 16 point provides a better user experience. As many previous studies investigated text with 10, 12, 14, and 16 point, there is an important question over the effect of 15 point. Does it create better user performance and do users prefer it in comparison with 14 and 16 point? Furthermore, it may be asked, does text larger than 16 point provide better performance and preference for older adults? These questions should be investigated in relation to all the text presentation variables investigated in this thesis: line spacing, font types, font size, and text/background colour combinations.

Extending the research to other aspects of web design

While the experiments in this thesis have analysed line spacing, text justification, font type, font size, and text colour and background colour combination, there are many other aspects of web design that should be studied to find empirical evidence for best practice for both younger and older users. Both evidence about user performance and preference measures are important to support recommendations and web design guidelines. Some important examples are line length, and the use of illustrations and animation on webpages. SilverWeb (2005, 2007) is the only set of guidelines to provide a recommendation on line length, but many participants comments on this variable. And the recommendation from SilverWeb is not clear as it suggests only that "text line should be short in length".

The different set of guidelines make different recommendations for illustrations and animation. For example, to provide realistic illustration or text-relevant images only (NIA/NLM, 2002; Silverweb, 2005, 2007; Zhao, 2001), to provide animation (SPRY Foundation, 1999; Holt, 2000), use short segments to avoid download time (NIA/NLM, 2002), provide animation but allow users to allow users to pause or stop (Zhao, 2001), animation can be distracting users (SPRY Foundation, 1999; Holt, 2000, AgeLight, 2001), animation should be avoid (SilverWeb, 2005, 2007).

Extending the study of web design guidelines in other languages and writing systems

Some recommendations from the current study for the UK participants are totally different from the recommendations for Thai participants, such as proposing sans serif for the UK participants and proposing serif for Thai participants. Some recommendations seem to depend on age group such as recommending black on white for older adults while recommending both black on white and sepia on off-white for younger adults. Finally, some recommendations are appropriate for all the user groups investigated, for example 1.5 or double line spacing are good for all groups of participants in the study. Thus, conducting further research on web design guidelines with participants using other languages and other writing systems will help in predicting which guidelines are age group specific, which recommendations might cover more than one language, and which ones cover more than one writing system.

Extending the research to reading the web on other devices

I found that older participants both in the UK and Thailand now have experience with reading the web with devices other than computer screens. Increasingly, older people are using tablet computers and smartphones to read webpages and other materials. In this thesis I was only able to conduct experiments about reading on a PC monitor, the results may well not generalize to reading from other devices. To establish guidelines for different devices, further research is necessary to inform the development of device-specific guidelines and to help older people reduce the numerous barriers in using the web because of age-related physical, sensory, and cognitive capabilities.

Extending the web design guidelines for Thai

While the current research programme has proposed the first web design guidelines relating to text presentation for the Thai alphabet for both younger and older adults, in future work I will investigate other guidelines for Thai older web users. However, it is also important to provide web accessibility and usability to Thai people with disabilities, because the web is an important medium which can provide opportunities and quality of life for people with disabilities. Therefore, further research is needed on guidelines about web accessibility for people with disabilities in the Thai language.

Appendix 1: Informed Consent Form (English)

Informed Consent Form

Thank you for participating in this study. This study is investigating the experience that people have when surfing the web with different types of line and page formatting.   At the beginning of the session you will be asked to complete a short questionnaire with demographic information. You will then be asked to undertake a series of short tasks to find information on a website. Finally, when all of the tasks are complete, you will be asked to complete another short questionnaire and will be given an opportunity to ask any questions you would like about this research study.

All information received during this study will be treated confidentially, and any results will be published in way that protects the anonymity of our participants. If you have any questions during the session please feel free to ask. Further, you may withdraw from the study at any time.

_____

Section A

I, _____, voluntarily consent to participate in this study on user experience of websites. I have been briefed about the basic nature and purpose of the project and feel that I understand it.

I understand that all data gathered will be treated confidentially. I understand that my data will only be available in its original form to Sorachai Kamollimsakul, Prof. Helen Petrie and Dr Christopher Power. I understand that I will not be identified when the data is shared, described or interpreted.

I also understand that I may withdraw at any point during the study.

_____   _____
Signature of research participant   Date

_____   _____
Signature of researcher   Date

Researcher contact details: Sorachai Kamollimsakul, sk750@york.ac.uk**,** +44 (0) 7886314499
_____

Section B

I have been adequately debriefed. I was not forced to complete the study. All my questions have been answered. I have been compensated for my participation as agreed.

Your signature: _____

Appendix 2: Informed Consent Form (Thai)

หนังสือแสดงเจตนายินยอมเข้าร่วมการวิจัย (Informed Consent Form)

ขอขอบคุณท่านที่เข้าร่วมในการศึกษาวิจัยนี้ การวิจัยนี้มุ่งสำรวจประสบการณ์ของผู้ใช้งานเว็บ เมื่อใช้งานเว็บที่มี
ประเภทของบรรทัดและรูปแบบการจัดหน้าที่แตกต่างกัน เมื่อเริ่มต้นการวิจัย ท่านจะถูกขอให้กรอก
แบบสอบถามสั้นๆเกี่ยวกับลักษณะทางประชากรศาสตร์ จากนั้นท่านจะถูกขอให้หาคำตอบของชุดคำถาม โดย
คำตอบหาได้จากข้อมูลที่ปรากฏในเว็บไซต์ และเมื่อท่านหาคำตอบของชุดตอบคำถามเสร็จสิ้น ท่านจะถูกขอให้
ตอบแบบสอบถามสั้นๆอีกชุดหนึ่ง จากนั้นท่านสามารถสอบถามรายละเอียดที่ท่านต้องการเกี่ยวกับการ
ศึกษาวิจัยนี้

ข้อมูลดิบทั้งหมดที่ได้รับระหว่างการศึกษาวิจัยนี้ถือเป็นความลับ ข้อมูลจะถูกเผยแพร่เป็นผลการวิจัยโดยรวม
โดยไม่ระบุตัวบุคคลแต่อย่างใด หากท่านมีคำถามระหว่างการวิจัยนี้ท่านสามารถสอบถามได้ตลอดเวลา และท่าน
สามารถถอนตัวจากการวิจัยได้ตลอดเวลา

_____

ส่วนที่ 1

ข้าพเจ้า _____ ขอแสดงเจตนายินยอมอย่างสมัครใจในการ
เข้าร่วมการศึกษาวิจัยเกี่ยวกับประสบการณ์ของผู้ใช้ที่มีต่อเว็บไซต์ ข้าพเจ้าได้รับทราบและเข้าใจรายละเอียด
เบื้องต้นและวัตถุประสงค์ของการวิจัยนี้

ข้าพเจ้ารับทราบว่าข้อมูลทั้งหมดจะเป็นความลับ ข้าพเจ้าเข้าใจดีว่ามีเพียง นายสรชัย กมลลิ้มสกุล
ศาสตราจารย์เฮเลน เพทรี (Prof. Helen Petrie) และ ดร.คริสโตเฟอร์ พาวเวอร์ (Dr. Christopher Power)
เท่านั้นที่จะเข้าถึงข้อมูลดิบของข้าพเจ้าที่ได้จากการวิจัย ข้าพเจ้ารับทราบว่าในการเผยแพร่ข้อมูล อธิบาย
ตีความจะไม่มีการระบุตัวตนของข้าพเจ้า

ข้าพเจ้ารับทราบว่าข้าพเจ้าสามารถถอนตัวจากการวิจัยได้ตลอดเวลา

_____   _____

ลายมือชื่อผู้เข้าร่วมการวิจัย   วัน เดือน ปี

_____   _____

ลายมือชื่อนักวิจัย   วันเดือนปี

รายละเอียดการติดต่อนักวิจัย: สรชัย กมลลิ้มสกุล อีเมล์: sk750@york.ac.uk โทรศัพท์: 09 - 5776 - 2595

_____

ส่วนที่ 2

ข้าพเจ้าถูกสอบถามรายละเอียดอย่างเพียงพอเหมาะสม และไม่ได้ถูกบังคับให้ทำการศึกษาวิจัยนี้ คำถามทุก
คำถามของข้าพเจ้าได้ถูกตอบ และข้าพเจ้าได้รับการตอบแทนในการเข้าร่วมการศึกษาวิจัยนี้ตามที่ตกลงไว้

ลายมือชื่อ: _____

Appendix 3: Pre-questionnaire

Initial Questions

This study is being conducted by Sorachai Kamollimsakul, a PhD student in the Human Computer Interaction Research Group, Department of Computer Science, University of York. The aim of the study is to investigate people's experience when using the web.

These initial questionnaires consist of 3 parts:

> **Part 1** Questions on your attitudes towards the Web

> **Part 2** Questions about yourself

> **Part 3** Questions on your use of the Web

Please answer all the questions, it should only take about 15 minutes to complete.

Researcher contact details: Sorachai Kamollimsakul, sk750@york.ac.uk**,** +44 (0) 7886314499

Part 1 Attitudes towards the Web

For each of the following statements, please choose the response on the rating scale which best indicates how much <u>you</u> agree or disagree with that statement. Please respond to every statement in terms of <u>your own experience</u> of using the Web.

If you are not completely sure which response to choose, put the response which you feel is most appropriate. Do not spend too long on each statement.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1. The Web is efficient	☐	☐	☐	☐	☐
2. The Web makes me nervous	☐	☐	☐	☐	☐
3. The Web has many useful features	☐	☐	☐	☐	☐
4. Fashionable people use the Web	☐	☐	☐	☐	☐
5. I have difficulty remembering how to use the Web	☐	☐	☐	☐	☐
6. The Web is reliable	☐	☐	☐	☐	☐
7. I feel anxious when using the Web	☐	☐	☐	☐	☐
8. The Web helps me to do a task effectively	☐	☐	☐	☐	☐
9. I can confidently operate the Web	☐	☐	☐	☐	☐
10. Learning to use the Web is easy	☐	☐	☐	☐	☐
11. Using the Web is good for my image	☐	☐	☐	☐	☐
12. I often need to refer to a manual for help	☐	☐	☐	☐	☐
13. The Web is rather difficult to use	☐	☐	☐	☐	☐
14. The Web is the best option for the job	☐	☐	☐	☐	☐

15. When I use the Web, I am afraid I will break/crash it	☐	☐	☐	☐	☐
16. The Web is value for money	☐	☐	☐	☐	☐
17. The Web does not make life easier	☐	☐	☐	☐	☐
18. I use the Web because lots of other people use it	☐	☐	☐	☐	☐

Part 2 Personal data

Finally, please answer the following general questions about yourself (this information is confidential and remember, we do not want your name)

1. What is your age?   .... Years

2. Are you?   Male ☐   Female ☐

3. Are you?
    - ☐ An undergraduate student
    - ☐ A Masters student
    - ☐ A Ph.D. student
    - ☐ Employed
    - ☐ Retired
    - ☐ Other (please specify).....................................

Part 3 Use of the Web

The following questions are about your use of the Web.

1. **For approximately how long have you been using the Web?**
   ☐ Less than 6 months ☐ 6-12 months ☐ 1-3 years
   ☐ 4-6 years   ☐ 7 years or more

2. **How did you learn to use the Web (tick all that applies)?**

   ☐ Self-taught ☐ Family and Relatives   ☐ friends / co-workers

   ☐ Courses   ☐ Work   ☐ Sale person

3. **How often do you use the Web per week?**

   ☐ Never   ☐ 1-5 hours   ☐ 6-10 hours
   ☐ 11-20 hours   ☐ more than 20 hours

4. **What is your level of computer experience (tick one)?**
   ☐ 1. None at all
   ☐ 2.
   ☐ 3.
   ☐ 4.
   ☐ 5.
   ☐ 6.
   ☐ 7. Extensive

5. **How expert do you feel about using the Web (tick one)?**
   ☐ 1. None at all
   ☐ 2.
   ☐ 3.
   ☐ 4.
   ☐ 5.
   ☐ 6.
   ☐ 7. Expert

Thank you for taking the time to complete these questions

Appendix 4: Pre-questionnaire (Thai)

<div align="center">

แบบสอบถามข้อมูลเบื้องต้น

</div>

แบบสอบถามนี้เป็นส่วนหนึ่งของการศึกษาของนายสรชัย กมลลิ้มสกุล นักศึกษาระดับปริญญาเอก กลุ่มวิจัยด้าน

ปฏิสัมพันธ์ระหว่างมนุษย์กับคอมพิวเตอร์ คณะวิทยาการคอมพิวเตอร์ มหาวิทยาลัยยอร์ก สหราชอาณาจักร

โดยมีวัตถุประสงค์เพื่อสำรวจประสบการณ์ของผู้ใช้เมื่อใช้งานเว็บ

แบบสอบถามข้อมูลเบื้องต้นนี้ประกอบด้วย 3 ส่วน

     ส่วนที่ 1 คำถามเกี่ยวกับทัศนคติที่มีต่อเว็บ

     ส่วนที่ 2 คำถามเกี่ยวกับข้อมูลส่วนตัวของท่าน

     ส่วนที่ 3 คำถามเกี่ยวกับการใช้งานเว็บ

กรุณาตอบแบบสอบถามให้ครบทุกข้อ การตอบแบบสอบถามใช้เวลาประมาณ 15 นาที

ติดต่อนักวิจัย : นายสรชัย กมลลิ้มสกุล sk750@york.ac.uk

ส่วนที่ 1 คำถามเกี่ยวกับทัศนคติที่มีต่อเว็บ (Attitudes towards the Web)

จากข้อความด้านล่างต่อไปนี้ ขอให้ท่านเลือกมาตรประมาณค่า (rating scale) ที่เหมาะสมที่สุดกับ ระดับความเห็นด้วย หรือไม่เห็นด้วยของ<u>ท่าน</u>ที่มีต่อข้อความนั้น ๆ กรุณาตอบคำถามทุกข้อความ บนพื้นฐานของ<u>ประสบการณ์ของท่าน</u>ในการใช้เว็บ

หากไม่มีมาตรประมาณค่าที่ตรงกับความคิดเห็นของท่าน ขอให้ท่านเลือกมาตรประมาณค่าที่ท่าน คิดว่าสอดคล้องมากที่สุด กรุณาอย่าใช้เวลานานในข้อความแต่ละข้อ

	ไม่เห็น ด้วย อย่างยิ่ง	ไม่เห็น ด้วย	เฉย ๆ	เห็น ด้วย	เห็นด้วย อย่างยิ่ง
1. เว็บมีประสิทธิภาพ	☐	☐	☐	☐	☐
2. เว็บทำให้ฉันรู้สึกกระวนกระวาย	☐	☐	☐	☐	☐
3. เว็บมีเครื่องมือที่ประโยชน์มากมาย	☐	☐	☐	☐	☐
4. คนที่ทันสมัยจะใช้งานเว็บ	☐	☐	☐	☐	☐
5. ฉันประสบปัญหาในการจำเกี่ยวกับการใช้งานเว็บ	☐	☐	☐	☐	☐
6. เว็บมีความน่าเชื่อถือ	☐	☐	☐	☐	☐
7. ฉันรู้สึกวิตกกังวลเมื่อใช้งานเว็บ	☐	☐	☐	☐	☐
8. เว็บช่วยให้ฉันทำงานได้อย่างได้ผล	☐	☐	☐	☐	☐
9. ฉันสามารถใช้งานเว็บได้อย่างมั่นใจ	☐	☐	☐	☐	☐
10. การเรียนรู้ที่จะใช้งานเว็บเป็นเรื่องง่าย	☐	☐	☐	☐	☐
11. การใช้งานเว็บดีต่อภาพลักษณ์ของฉัน	☐	☐	☐	☐	☐
12. บ่อยครั้งที่ฉันต้องการคู่มือเพื่อช่วยในการใช้งาน	☐	☐	☐	☐	☐
13. การใช้งานเว็บเป็นเรื่องค่อนข้างยาก	☐	☐	☐	☐	☐
14. เว็บเป็นทางเลือกที่ดีที่สุดสำหรับการทำงาน	☐	☐	☐	☐	☐

15. เมื่อฉันใช้งานเว็บ ฉันกลัวที่จะทำพัง	☐	☐	☐	☐	☐
16. เว็บมีคุณค่าสมกับเงินที่จ่ายไป	☐	☐	☐	☐	☐
17. เว็บไม่ช่วยให้ชีวิตง่ายขึ้น	☐	☐	☐	☐	☐
18. ฉันใช้เว็บเพราะคนจำนวนมากใช้มัน	☐	☐	☐	☐	☐

ส่วนที่ 2 คำถามเกี่ยวกับข้อมูลส่วนตัวของท่าน

กรุณาตอบคำถามเกี่ยวกับข้อมูลส่วนตัวของท่าน (ข้อมูลนี้จะเป็นความลับ และท่านไม่ต้องระบุชื่อ-สกุลของท่าน)

1. ท่านอายุ   .... ปี

2. เพศ   ☐ ชาย   ☐ หญิง

3. อาชีพ

        ☐ นักศึกษาระดับปริญญาตรี
        ☐ นักศึกษาระดับปริญญาโท
        ☐ นักศึกษาระดับปริญญาเอก
        ☐ ทำงาน
        ☐ เกษียณอายุ
        ☐ อื่นๆ (โปรดระบุ)....................................

ส่วนที่ 3 คำถามเกี่ยวกับการใช้งานเว็บ

คำถามต่อไปนี้เกี่ยวกับการใช้งานเว็บของท่าน

1. ท่านใช้งานเว็บมานานเท่าไร?

☐ น้อยกว่า 6 เดือน   ☐ 6-12 เดือน   ☐ 1 - 3 ปี

☐ 4 - 6 ปี   ☐ 7 ปี หรือมากกว่า

2. ท่านเรียนรู้การใช้งานเว็บได้อย่างไร? (เลือกคำตอบได้มากกว่า 1 ข้อ)

☐ เรียนด้วยตนเอง   ☐ จากคนในครอบครัว หรือญาติ   ☐ จากเพื่อน หรือเพื่อนร่วมงาน

☐ หลักสูตรการอบรม   ☐ จากที่ทำงาน   ☐ จากพนักงานขาย

3. ท่านใช้งานเว็บเป็นระยะเวลาเท่าใดใน 1 สัปดาห์?

☐ ไม่เคยใช้เลย   ☐ 1-5 ชั่วโมง   ☐ 6-10 ชั่วโมง

☐ 11-20 ชั่วโมง   ☐ มากกว่า 20 ชั่วโมง

4. ประสบการณ์การใช้คอมพิวเตอร์ของท่านอยู่ในระดับใด? (กรุณาเลือกเพียงข้อเดียว)

☐ 1. ไม่มีประสบการณ์เลย

☐ 2.

☐ 3.

☐ 4.

☐ 5.

☐ 6.

☐ 7. มีประสบการณ์อย่างกว้างขวาง

5. ท่านรู้สึกว่าตัวท่านมีความเชี่ยวชาญในการใช้เว็บอยู่ในระดับใด? (กรุณาเลือกเพียงข้อเดียว)

☐ 1. ไม่มีความเชี่ยวชาญเลย

☐ 2.

☐ 3.

☐ 4.

☐ 5.

☐ 6.

☐ 7. มีความเชี่ยวชาญเป็นอย่างมาก

ขอขอบคุณท่านที่สละเวลาในการตอบคำถาม

Appendix 5: Post-questionnaire (English)

Post-study questions

1. In the website you experienced three line spacings.

1(a) For each line spacing, please rate how **easy** it was to read the text with that line spacing.
[use the sheet provided to remind yourself of the spacings]

	Very easy	Quite easy	Neither easy or hard	Quite difficult	Very difficult
Single spacing					
1.5 spacing					
Double spacing					

1(b) For each line spacing, please rate how **pleasant** it was to read the text with that line spacing. [use the sheet provided to remind yourself of the spacings]

	Very unpleasant	Quite unpleasant	Neutral	Quite pleasant	Very pleasant
Single spacing					
1.5 spacing					
Double spacing					

1(c) For each line spacing, please rate how **fast** it was to read the text with that line spacing.
[use the sheet provided to remind yourself of the spacings]

	Very slow	Quite slow	Average	Quite fast	Very fast
Single spacing					
1.5 spacing					
Double spacing					

2. In the website you experienced two justifications.

2 (a) For each line justification, please rate how **easy** it was to read the text with that line justification. [use the sheet provided to remind yourself of the justifications]

	Very easy	Quite easy	Neither easy or hard	Quite difficult	Very difficult
Left justified					
Left – right justified					

2 (b) For each line justification, please rate how **pleasant** it was to read the text with that line justification. [use the sheet provided to remind yourself of the justifications]

	Very unpleasant	Quite unpleasant	Neutral	Quite pleasant	Very pleasant
Left justified					
Left – right justified					

2 (c) For each line justification, please rate how **fast** it was to read the text with that line justification. [use the sheet provided to remind yourself of the justifications]

	Very slow	Quite slow	Average	Quite fast	Very fast
Left justified					
Left – right justified					

3. Which overall combination of line-spacing and line justification did you prefer most? [use the sheet provided to remind yourself of the combinations]

	Least prefer	Less prefer	Neutral	Quite prefer	Most prefer
Single line spacing with left justified					
1.5 line spacing with left justified					
double lines spacing with left justified					
Single line spacing with left-right justified					
1.5 lines spacing with left-right justified					
Double lines spacing with left-right justified					

Could you explain why you most prefer that combination?

....
....
....
....
....

4. Are there other aspects of the presentation of text on websites that particularly irritate you – for example, font size, line length, font type etc.

....
....
....
....
....
....

Thank you for taking the time to complete these questions

Appendix 6: Post-questionnaire (Thai)

แบบสอบถามหลังการศึกษา

1. ในเว็บไซต์ ท่านจะพบระยะห่างระหว่างบรรทัด 3 รูปแบบ

1(a) กรุณาประเมินว่าระยะห่างระหว่างบรรทัดแต่ละแบบทำให้เกิด **ความง่าย** (ease) ต่อการอ่านข้อความบน
หน้าเว็บอย่างไรบ้าง [ท่านสามารถดูตัวอย่างระยะห่างระหว่างบรรทัดที่แตกต่างกันเพื่อประกอบการประเมิน]

	ง่ายมาก	ค่อนข้างง่าย	ไม่ง่าย และ ไม่ยาก	ค่อนข้างยาก	ยากมาก
ระยะห่าง 1 บรรทัด					
ระยะห่าง 1.5 บรรทัด					
ระยะห่าง 2 บรรทัด					

1(b) กรุณาประเมินว่าระยะห่างระหว่างบรรทัดแต่ละแบบทำให้เกิด **ความน่าพอใจ** (pleasantness) ต่อการอ่าน
ข้อความบนหน้าเว็บอย่างไรบ้าง [ท่านสามารถดูตัวอย่างระยะห่างระหว่างบรรทัดที่แตกต่างกันเพื่อประกอบการ
ประเมิน]

	ไม่พึงพอใจ อย่างมาก	ไม่ค่อย พึงพอใจ	เฉยๆ	ค่อนข้าง พึงพอใจ	พึงพอใจ อย่างมาก
ระยะห่าง 1 บรรทัด					
ระยะห่าง 1.5 บรรทัด					
ระยะห่าง 2 บรรทัด					

1(c) กรุณาประเมินว่าระยะห่างระหว่างบรรทัดแต่ละแบบทำให้เกิด **ความเร็ว** (speed) ต่อการอ่านข้อความบน
หน้าเว็บอย่างไรบ้าง [ท่านสามารถดูตัวอย่างระยะห่างระหว่างบรรทัดที่แตกต่างกันเพื่อประกอบการประเมิน]

	ช้ามาก	ค่อนข้างช้า	ปานกลาง	ค่อนข้างเร็ว	เร็วมาก
ระยะห่าง 1 บรรทัด					
ระยะห่าง 1.5 บรรทัด					
ระยะห่าง 2 บรรทัด					

2. ในเว็บไซต์ ท่านจะพบการจัดแนวข้อความ 2 รูปแบบ

2 (a) กรุณาประเมินว่าการจัดแนวข้อความแต่ละแบบทำให้เกิด **ความง่าย** (ease) ต่อการอ่านข้อความบนหน้า เว็บอย่างไรบ้าง [ท่านสามารถดูตัวอย่างการจัดแนวข้อความที่แตกต่างกันเพื่อประกอบการประเมิน]

	ง่ายมาก	ค่อนข้างง่าย	ไม่ง่ายและไม่ยาก	ค่อนข้างยาก	ยากมาก
การจัดแนวข้อความแบบ ชิดซ้าย					
การจัดแนวข้อความแบบ เสมอหน้าหลัง					

2 (b) กรุณาประเมินว่าการจัดแนวข้อความแต่ละแบบทำให้เกิด **ความน่าพอใจ** (pleasantness) ต่อการอ่าน ข้อความบนหน้าเว็บอย่างไรบ้าง [ท่านสามารถดูตัวอย่างการจัดแนวข้อความที่แตกต่างกันเพื่อประกอบการ ประเมิน]

	ไม่พึงพอใจอย่างมาก	ไม่ค่อยพึงพอใจ	เฉยๆ	ค่อนข้างพึงพอใจ	พึงพอใจอย่างมาก
การจัดแนวข้อความแบบ ชิดซ้าย					
การจัดแนวข้อความแบบ เสมอหน้าหลัง					

2 (c) กรุณาประเมินว่าการจัดแนวข้อความแต่ละแบบทำให้เกิด **ความเร็ว** (speed) ต่อการอ่านข้อความบนหน้า เว็บอย่างไรบ้าง [ท่านสามารถดูตัวอย่างการจัดแนวข้อความที่แตกต่างกันเพื่อประกอบการประเมิน]

	ช้ามาก	ค่อนข้างช้า	ปานกลาง	ค่อนข้างเร็ว	เร็วมาก
การจัดแนวข้อความแบบ ชิดซ้าย					
การจัดแนวข้อความแบบ เสมอหน้าหลัง					

3. การผสมระหว่างระยะห่างระหว่างบรรทัด และการจัดแนวข้อความ แบบใดที่ท่านชื่นชอบมากที่สุด? [ท่าน
สามารถดูตัวอย่างระยะห่างระหว่างบรรทัด และการจัดแนวข้อความที่แตกต่างกันเพื่อประกอบการประเมิน]

	ชื่นชอบ น้อยที่สุด	ชื่นชอบ น้อย	เฉยๆ	ชื่นชอบ มาก	ชื่นชอบ มากที่สุด
ระยะห่าง 1 บรรทัด กับ การจัดแนวข้อความแบบชิดซ้าย					
ระยะห่าง 1.5 บรรทัด กับ การจัดแนวข้อความแบบชิดซ้าย					
ระยะห่าง 2 บรรทัด กับ การจัดแนวข้อความแบบชิดซ้าย					
ระยะห่าง 1 บรรทัด กับ การจัดแนวข้อความแบบเสมอหน้าหลัง					
ระยะห่าง 1.5 บรรทัด กับ การจัดแนวข้อความแบบเสมอหน้าหลัง					
ระยะห่าง 2 บรรทัด กับ การจัดแนวข้อความแบบเสมอหน้าหลัง					

โปรดอธิบายว่าเพราะเหตุใดท่านจึงชื่นชอบการผสมระหว่างระยะห่างระหว่างบรรทัด และการจัดแนวข้อความ
แบบนั้นมากที่สุด

....
....
....
....
....
....

4. มีลักษณะการนำเสนอข้อความในเว็บไซต์อื่นใดอีกหรือไม่ที่ส่งผลต่อท่าน เช่น ขนาดตัวอักษร ความยาวของ
บรรทัด ชนิดของตัวอักษร ฯลฯ

....
....
....
....
....

Appendix 7: Olympic Games web site (English and Thai version)

The Olympic Game web site in English version:

- Single line spacing with left justification condition:

http://yorkhci.org/olympic1/?q=node/131

- 1.5 line spacing with left justification condition:

http://yorkhci.org/olympic2/?q=node/131

- Double line spacing with left justification condition:

http://yorkhci.org/olympic3/?q=node/131

- Single line spacing with left- right justification condition:

http://yorkhci.org/olympic4/?q=node/131

- 1.5 line spacing with left- right justification condition:

http://yorkhci.org/olympic5/?q=node/131

- Double line spacing with left- right justification condition:

http://yorkhci.org/olympic6/?q=node/131

The Olympic Game web site in Thai version:

- Single line spacing with left justification condition:

https://dl.dropboxusercontent.com/u/23143888/Web_Thai_Final/olympic_t1/index.html

- 1.5 line spacing with left justification condition:

https://dl.dropboxusercontent.com/u/23143888/Web_Thai_Final/olympic_t2/index.html

- Double line spacing with left justification condition:

https://dl.dropboxusercontent.com/u/23143888/Web_Thai_Final/olympic_t3/index.html

- Single line spacing with left- right justification condition:

https://dl.dropboxusercontent.com/u/23143888/Web_Thai_Final/olympic_t4/index.html

- 1.5 line spacing with left- right justification condition:

https://dl.dropboxusercontent.com/u/23143888/Web_Thai_Final/olympic_t5/index.html

- Double line spacing with left- right justification condition:

https://dl.dropboxusercontent.com/u/23143888/Web_Thai_Final/olympic_t6/index.html

Appendix 8: Tasks in experiment 1 (English version)

1. How many times has the Olympic flame been carried across the water?

Ans. ....

2. In the first modern Olympic games of 1896, who was considered to be the most successful athlete of those games?

Ans. ....

3. At Beijing in 2008, how many years was it since the UK had won a gold medal in swimming?

Ans. ....

4. Who was the first woman from Thailand to earn a gold medal in the 2004 Athens Olympic Games?

Ans. ....

5. What are the common reasons for a substitution of a player during a football match?

Ans. ....

6. What is the maximum width of baselines in tennis court?

Ans. ....

Thank you for taking the time to complete these tasks

These six tasks were divided between the four main web pages as shown in below;

About the games	1	task
History of the games	1	task
Previous Olympic Games	2	tasks
Olympic Sports	2	tasks

Figure 1 shows that "About the games" was categorised into 6 pages; International Olympic Committee (IOC), Olympic traditions, Olympic motto, Olympic rings, Opening ceremonies, and Closing ceremonies. One task asked a question about the Olympic flame which the answer can be found in "The Torches" under "Olympic traditions".

Task 1: How many times has the Olympic flame been carried across the water?

Answer: 2 times

Optimal Path: Home > About the Games > Olympic Traditions > Torches
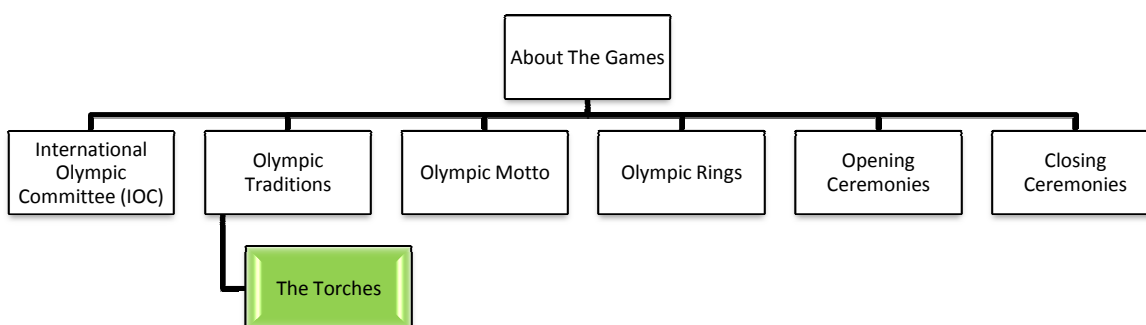
Figure 1: Web pages under the content about "About the Games" with a web page where task1's answer was placed

Figure 2, shows that "History of the games" was separated into 2 pages; ancient Olympic Games, and modern Olympic Games. Ancient Olympic Games was separated into 4 sub-pages; About history, Methodology, The athlete, and The sport events. Modern Olympic Games was separated into 4 pages; Forerunners, Revival, 1896 games, and Expanded, changed, and adapted. The answer of the task in this part is in "1896 Games".

Task 2: In the first modern Olympic games of 1896, who was considered to be the most successful athlete of those games?

Answer: Carl Schumann

Optimal Path: Home › History of the Games › Modern Olympics › 1896 Games
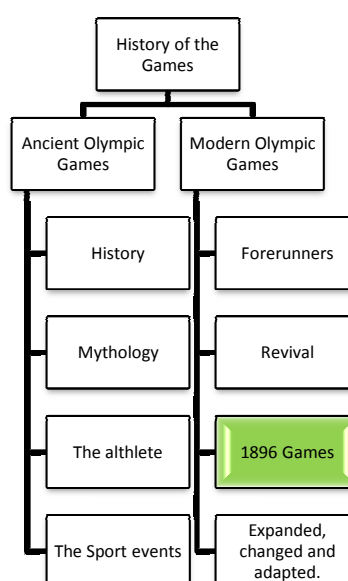
Figure 2: Web pages under the content about "History of the Games" with a web page where task2's answer was placed

In previous Olympic Games, there were information about Olympics events in the past since the first modern Olympic Games, Athens 1896 to Beijing 2008. However, there were only 2 pages which had hypertext links lead to information about two Olympic Games events; Beijing 2008 and Athens 2004. One task asked about information related to the UK, which an answer appeared in "interesting facts", in Beijing 2008. Another task asked about a Thai athlete. An answer in this task was on "interesting facts", in Athens 2004 as shown in Figure 3.

Task 3: At Beijing in 2008, how many years was it since the UK had won a gold medal in swimming?

Answer: Almost 50 years (48 years)

Optimal Path: Home › Previous Olympic Games › Beijing 2008 › Beijing 2008 in detail › Interesting Facts

Task 4: Who was the first woman from Thailand to earn a gold medal in the 2004 Athens Olympic Games?

Answer: Pawina Thongsuk

Optimal Path: Home › Past Olympic Games › Athens 2004 › Athens 2004 in detail › Interesting Facts
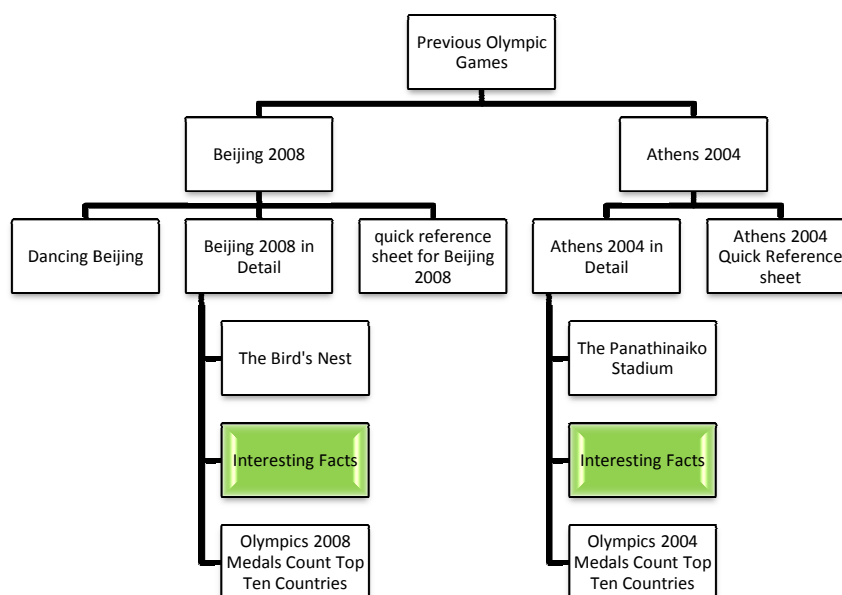
Figure 3: Web pages under the content about "Previous Olympic Games" with a web page where task3 and task4's answers were placed

In Olympic sports, there is information about 7 sports in 7 pages. There are Football, Aquatics, Athletics, Badminton, Tennis, Basketball, and Gymnastics. Two tasks are in Olympic sports; one task asks about players and officials in Football and another task asks about the field of play in Tennis as shown in Figure 4.

Task 5: What are the common reasons for a substitution of a player during a football match?

Answer: injury, tiredness, ineffectiveness, a tactical switch, or time wasting

Optimal Path: Home > Olympic Sports > Football > Rules and gameplay > Players and Officials

Task 6: What is the maximum width of baselines in tennis court?

Answer: The baseline can be up to 4 inches (100 mm) wide if so desired.

Optimal Path: Home › Olympic Sports › Tennis › Manner of Play › Field of Play
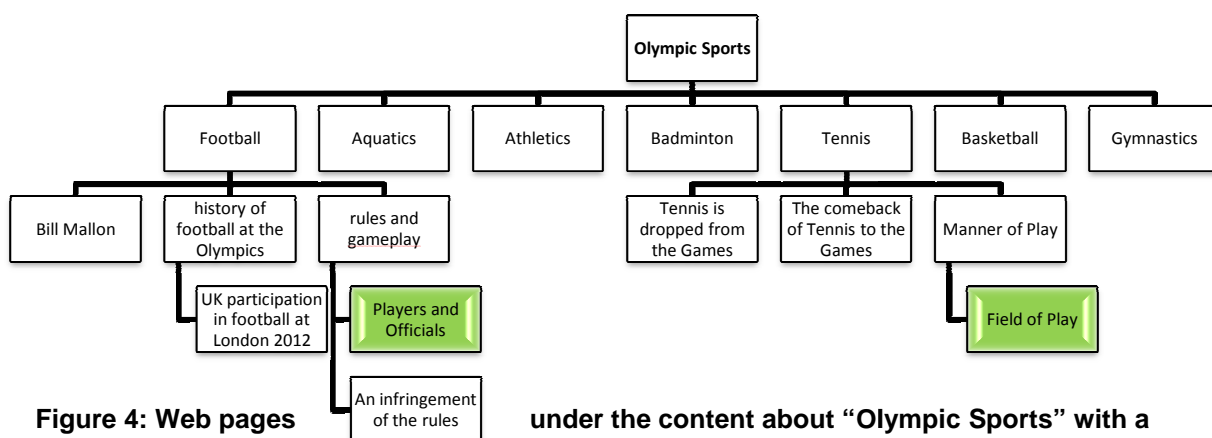
Figure 4: Web pages under the content about "Olympic Sports" with a web page where task5 and task6's answers were placed

Appendix 9: Tasks in experiment 1 (Thai version)

1. ในการแข่งขันโอลิมปิคเกมส์ ไฟโอลิมปิคถูกนำข้ามน้ำทั้งหมดกี่ครั้ง?

   คำตอบ...

2. ในการแข่งขันโอลิมปิกสมัยใหม่ครั้งแรกที่จัดขึ้นเมื่อปี ค.ศ.1896 ใครคือนักกีฬาที่ประสบความสำเร็จ
   มากที่สุด?

   คำตอบ...

3. ในการแข่งขันกีฬาโอลิมปิค ปักกิ่ง 2008 อังกฤษได้เหรียญทองจากกีฬาว่ายน้ำอีกครั้ง หลังจากที่เคย
   ทำได้เมื่อกี่ปีมาแล้ว?

   คำตอบ...

4. ใครคือนักกีฬาหญิงชาวไทยคนแรกที่ได้เหรียญทอง จากการแข่งขันกีฬาโอลิมปิค เอเธนส์ 2004?

   คำตอบ...

5. การเปลี่ยนตัวผู้เล่นระหว่างการแข่งขันกีฬาฟุตบอล มีสาเหตุอะไรบ้าง?

   คำตอบ...

6. เส้นหลังของสนามเทนนิส มีความกว้างได้มากที่สุดเท่าใด?

   คำตอบ...

ขอขอบคุณท่านที่สละเวลาในการตอบคำถาม

Appendix 10: Attitude toward the Web Scoring sheet and actual factors

	Actual Factors	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1. The Web is efficient	P	1	2	3	4	5
2. The Web makes me nervous	C	5	4	3	2	1
3. The Web has many useful features	P	1	2	3	4	5
4. Fashionable people use the Web	F	1	2	3	4	5
5. I have difficulty remembering how to use the Web	C	5	4	3	2	1
6. The Web is reliable	P	1	2	3	4	5
7. I feel anxious when using the Web	C	5	4	3	2	1
8. The Web helps me to do a task effectively	P	1	2	3	4	5
9. I can confidently operate the Web	C	1	2	3	4	5
10. Learning to use the Web is easy	C	1	2	3	4	5
11. Using the Web is good for my image	F	1	2	3	4	5
12. I often need to refer to a manual for help	C	5	4	3	2	1
13. The Web is rather difficult to use	C	5	4	3	2	1
14. The Web is the best option for the job	C	1	2	3	4	5

15. When I use the Web, I am afraid I will break/crash it	C	5	4	3	2	1
16. The Web is value for money	P	1	2	3	4	5
17. The Web does not make life easier	P	5	4	3	2	1
18. I use the Web because lots of other people use it	F	1	2	3	4	5

(P = Performance; C= Confidence; F = Fashion) (Burn, 2003 pp. 178, 344-345)

Appendix 11: Content, target words, and multiple choices questions for experiment

 2 - 4

1. (For practice trial in scanning task)

In 1968 – ironically the year in which tennis faced up to the facts of commercial life by accepting the concept of 'open' rather than strictly amateur events – tennis was included in the Olympic Games in Mexico, although only as a exhibition/demonstration sport.

 Tennis staged a 21 & under demonstration event at Los Angeles 1984, although by then the long, determined campaign to have tennis welcomed back as a full medal sport was well into its stride.

 The champion of the cause was David Gray, then General Secretary of the International Tennis Federation (ITF), who sadly died before all his work had come to fruition. His belief in the merits of tennis returning to the Olympic fold was unshakeable, and he had equally enthusiastic and determined support in this belief from the ITF President, Philippe Chatrier of France, and the Vice President, Pablo Llorens of Spain.

The Olympic Tennis Event in Los Angeles attracted capacity 6,000 crowds each day. Its success, as well as the growing awareness both within and beyond the IOC that Olympic **membership** assists with the grass roots development of any sport, made the decision to readmit tennis into the Olympics seem appropriate. It was ultimately decided that the world's finest tennis players should once again be allowed to compete for gold medals, along with their leading counterparts in other sports at this greatest of all sports gatherings

The respective singles winners in 1984, Steffi Graf led the way again as top seeds in Seoul. Graf went on to complete what has become known as her "Golden Slam" (she had already won all four Grand Slam tournaments that year).

Target word for scanning task: membership

2. For practice trial in skimming task

The Olympic Games are a major international event featuring summer and winter sports, in which thousands of athletes participate in a variety of competitions. The Olympic Games have come to be regarded as the world's foremost sports competition where more than 200 nations participate. The Games are currently held every four years within their respective seasonal games.

The modern games began in 1894 with Baron Pierre de Coubertin founding the International Olympic Committee (IOC). The IOC has since become the governing body of the Olympic Movement, its structure and actions are defined by the Olympic Charter. The Olympic Movement consists of International Sports Federations (IFs), National Olympic Committees (NOCs), and organizing committees for each specific Olympic Games. As the decision-making body, the IOC is responsible for choosing the host city for each Olympic Games. The host city is responsible for organizing and funding a celebration of the Games consistent with the Olympic Charter. The Olympic programme, consisting of the sports to be contested at the Games, is also determined by the IOC.

The celebration of the Games encompasses many rituals and symbols, such as the Olympic flag and torch, as well as the opening and closing ceremonies. Over 13,000 athletes compete at the Summer and Winter Olympics in 33 different sports and in nearly 400 events. The first, second, and third placed competitors in each event receive Olympic medals: gold, silver, and bronze, respectively.

The evolution of the Olympic Movement during the 20th and 21st centuries has resulted in several changes to the Olympic Games. Some of these adjustments include the creation of the Winter Games, the Paralympic Games, and the Youth Olympic Games.

Question for skimming task

1. Which of the following titles best fit the text?
   (a) Olympic Games
   (b) Summer Olympics Games
   (c) Olympic movement
   (d) I'm not sure
2. Approximately how many countries participate in the Olympic Games
   (a) More than 100 nations
   (b) More than 150 nations
   (c) More than 200 nations
   (d) I'm not sure
3. Who is Baron Pierre de Coubertin?
   (a) Founder of the International Olympic Committee (IOC)
   (b) Founder of the International Sports Federations (IFs)
   (c) Founder of the National Olympic Committees (NOCs)

(d) I'm not sure
4. Which does the evolution of the Olympic movement during the 20th and 21st centuries?
(a) Winter Games
(b) Winter Game and Paralympic Games
(c) Winter Games, Paralympic Game, and Youth Olympic Games
(d) I'm not sure

3. For practice trial in detailed reading task

The Federation of International Gymnastics (FIG) was founded in Liege in 1881. By the end of the nineteenth century, men's gymnastics competition was popular enough to be included in the first modern Olympic Games in 1896. Until the early 1950s, both national and international competitions involved a changing variety of exercises gathered under the rubric of gymnastics that would seem strange to today's audiences. These included synchronized team floor calisthenics, rope climbing, high jumping, running, and horizontal ladder. During the 1920s, women organized and participated in gymnastics events. The first women's Olympic competition was primitive, as it involved only synchronized calisthenics. It was held at the 1928 Games, in Amsterdam.

By 1954, apparatus and events for gymnastics at the Olympics Games had been standardized in their modern format for both men and women, and uniform grading structures (including a point system from 1 to 15) had been agreed upon. At this time, Soviet gymnasts astounded the world with highly disciplined and difficult performances, setting a precedent that continues.

The new medium of television helped publicize and initiate a modern age of gymnastics. Both men's and women's gymnastics now attract considerable international interest, and excellent gymnasts can be found on every continent. Nadia Comăneci received the first perfect score, at the 1976 Summer Olympics held in Montreal, Canada. She was coached in Romania by Béla Károlyi, a Romanian coach of Hungarian origin. Comăneci scored four of her perfect tens on the uneven bars, two on the balance beam and one in the floor exercise. Even with Nadia's perfect scores, the Romanians lost the gold medal to the Soviet Union. Nevertheless, Comăneci became an Olympic icon.

Question for detailed reading task

1. Which of the following titles best fit the text?
(a) History of Gymnastics
(b) The most successful gymnasts
(c) Effect of Television in modern age of Gymnastics
(d) I'm not sure

2. When does women organized and participated in gymnastics events?
   (a) 1986
   (b) 1920s
   (c) 1950s
   (d) I'm not sure

3. By 1954, which country astounded the world with highly disciplined and difficult performances?
   (a) United States
   (b) Soviet Union
   (c) Romania
   (d) I'm not sure

4. Who was the gymnastic Olympic icon at the Olympic Games in 1976 ?
   (a) Nadia Comăneci
   (b) Béla Károlyi
   (c) Soviet gymnasts
   (d) I'm not sure

Text 1

Athletic contests in running, walking, jumping and throwing are among the oldest of all sports and their roots are prehistoric. Athletics events were depicted on ancient Egyptian tombs, with illustrations of running and high jumping appearing from as early as 2250 BC. Similarly, the Tailteann Games were an ancient Celtic festival in Ireland which has been dated to have occurred around 1800 BC. This thirty-day meeting included running and stone-throwing among its sporting events.

The original and only athletics event at the first Olympics in 776 BC was a stadium-length running event known as the stadion. This was later expanded to be the ancient pentathlon which included throwing and jumping events. Athletics competitions also took place at other Panhellenic Games which were founded later, in around 500 BC.

The Cotswold Olimpick Games, a sports festival which emerged in 17th century England, also featured athletics in the form of a sledgehammer throwing contest. Similarly, an annual event called L'Olympiade de la République was held in revolutionary France from 1796 to 1798. This event was an early forerunner to the modern summer Olympic Games. The premier event of this competition was a running event, but various ancient Greek disciplines were also on display. The 1796 Olympiade marks the introduction of the metric system into sport.

The Modern Era of Athletics is attributed to have started at The Royal Military College, Sandhurst. This is claimed to be the first to hold athletics meetings in 1812 and 1825; but, there is no formal supporting **evidence** of this claim. The first modern-style indoor athletics meetings were recorded shortly after in the 1860s, including a meet at Ashburnham Hall in London.

Target word for scanning task: evidence

Question for skimming and detailed reading tasks

1. Where is the oldest historical record we have of athletics events?
   (a) Greece
   (b) France
   (c) Egypt
   (d) I'm not sure

2. What was the original and only athletics event at the first Olympics?
   (a) stadium-length running
   (b) high jumping
   (c) throwing
   (d) I'm not sure

3. When was the metric system introduced into sport?
   (a) 500 BC in Greece
   (b) 1796 in France
   (c) 1812 in England
   (d) I'm not sure

4. When was the first recorded indoor athletics meeting?
   (a) In the 1700s
   (b) In the 1790s
   (c) In the 1860s
   (d) I'm not sure

Text 2

The symbol of the Olympic Games is composed of five interlocking rings, coloured blue, yellow, black, green, and red on a white field. This was originally designed in 1912 by Baron Pierre de Coubertin, the founder of the modern Olympic Games. Upon its initial introduction, de Coubertin stated the following in the August, 1912 edition of Olympique:

The emblem chosen to illustrate and represent the World Congress of 1914...: five intertwined rings in different colours - blue, yellow, black, green, and red - are placed on the white field of the paper. These five rings represent the five parts of the world which now are won over to Olympism and willing to accept healthy competition.

In his article published in the "Olympic Revue" the official magazine of the International Olympic Committee in November 1992, the American historian Robert Barney explains that the idea of the interlaced rings came to Pierre de Coubertin when he was in charge of the USFSA, an association founded by the union of two French sports associations and until 1925, responsible for representing the International Olympic Committee in France. The emblem of the union was two interlaced rings and was originally the idea of Swiss psychiatrist Carl Jung. For Jung the ring meant **continuity** and the human being.

According to De Coubertin, the ring colours with the white background stand for those colours that appeared on all the national flags of the world at that time.

The 1914 Congress had to be suspended because of the outbreak of World War I, but the symbol and flag were later adopted. They would first officially debut at the Games of the VII Olympiad in Antwerp, Belgium in 1920.

Target word for scanning task: continuity

Question for skimming and detailed reading tasks

1. Who was the designer of the modern Olympic symbol of 5 coloured rings on a white background?
   (a) Baron Pierre de Coubertin
   (b) Robert Barney
   (c) Carl Jung
   (d) I'm not sure

2. What is the name of the official magazine of the International Olympic Committee?
   (a) IOC Magazine

      (b) Olympic Magazine
      (c) Olympic Revue
      (d) I'm not sure

3. For Swiss psychiatrist Carl Jung, what was the meaning of the Olympic rings?
      (a) Continuity and Human being
      (b) Olympic spirit and human being
      (c) Continuity and Olympic spirit
      (d) I'm not sure

4. When was the official debut of the Olympic rings?
      (a) 1912
      (b) 1914
      (c) 1920
      (d) I'm not sure

Text 3

Football was not on the programme at the first modern Olympic Games in 1896, as international soccer was in its infancy at the time. However, some sources claim that an unofficial football tournament was organized during the first competition, in which an Athens XI lost to a team representing Smyrna (Izmir), then part of the Ottoman Empire. Smyrna went on to be undefeated (15–0) by a team from Denmark. However, it is in fact unclear whether any competition took place at all; the Olympic historian Bill Mallon has written: "Supposedly a match between a Greek club and a Danish club took place. No such 1896 source supports this and we think this is an error which has been perpetuated in multiple texts. No such match occurred".

Tournaments were played at the 1900 and 1904 games and the Intercalated Games of 1906, but these were contested by various clubs and scratch teams. Although the IOC considers the 1900 and 1904 tournaments to be official Olympic events, they are not recognized by the International Federation of Association Football (FIFA); neither recognizes the Intercalated Games today. In 1906, teams from Great Britain, Germany, Austria, the Netherlands and France were withdrawn from an unofficial competition and left Denmark, Smyrna, Athens and Thessaloniki Music Club to compete. Denmark won the final against Athens 9–0.

After the initial tournament played in 1908 at the London Games, football became an important piece of the Olympic Games. The competition became increasingly important throughout the 1920s, although that decade witnessed a bad day in the history of the Olympic **movement** when, during the 1920 final, Czechoslovakia walked from the field of play.

Target word for scanning task: movement

Question for skimming and detailed reading tasks

1. Why was football not an official sport programme at Olympic Games in 1896?
   (a) International soccer was in its infancy
   (b) There was confusion about whether country or city teams could enter
   (c) Not enough countries played football at the time
   (d) I'm not sure

2. Which association or associations acknowledge the tournaments held in 1900 and 1904 as being part of the official Olympic Games?
   (a) The International Olympic Committee (IOC)
   (b) The International Federation of Association Football (FIFA)
   (c) Both the International Olympic Committee (IOC) and the International Federation of Association Football (FIFA)
   (d) I'm not sure

3. In the 1906 games, what team did Denmark defeat in the final?
   (a) Great Britain
   (b) Athens
   (c) France
   (d) I'm not sure

4. When did the Czechoslovakian team walk from the field of play in the football gold medal match?
   (a) 1904
   (b) 1908
   (c) 1920
   (d) I'm not sure

Text 4

As mandated by the Olympic Charter, various elements frame the opening ceremony of the Olympic Games. Most of these rituals were established at the 1920 Summer Olympics in Antwerp. The ceremony typically starts with the hoisting of the host country's flag and a performance of its national anthem. The host nation then presents artistic displays of music, singing, dance, and theatre representative of its culture.

After the artistic portion of the ceremony, the athletes parade into the stadium grouped by nation. Greece is traditionally the first nation to enter in order to honour the origins of the Olympics. Nations then enter the stadium alphabetically according to the host country's chosen language, with the host country's athletes being the last to enter. Speeches are given, formally opening the Games. Finally, the Olympic torch is brought into the stadium and passed on until it reaches the final torch carrier - often a well-known and successful Olympic athlete from the host nation - who lights the Olympic flame in the stadium's cauldron.

The closing ceremony of the Olympic Games takes place after all sporting events have concluded. Flag-bearers from each participating country enter the stadium, followed by the

athletes who enter together, without any national distinction. Three national flags are hoisted while the corresponding national anthems are played: the flag of Greece, to honour the **birthplace** of the Olympic Games; the flag of the current host country; and the flag of the country hosting the next Olympic Games.

The president of the organizing committee and the IOC president make their closing speeches, the Games are officially closed, and the Olympic flame is extinguished.

Target word for scanning task: birthplace

Question for skimming and detailed reading tasks

1. When were the rituals of the opening ceremony established?
   (a) 1896
   (b) 1912
   (c) 1920
   (d) I'm not sure

2. When are the Olympic Games officially open?
   (a) After opening speeches are given
   (b) After the Olympic flame is lit in the stadium cauldron.
   (c) After the hoisting of the Olympic flag
   (d) I'm not sure

3. Which three flags are hoisted while the corresponding national anthems are played in the closing ceremony?
   (a) The Greek flag, the flag of current host country, the Olympic flag
   (b) The Olympic flag, the flag of current host country, the flag of the next host country
   (c) The Greek flag, the flag of current host country, the flag of the next host country
   (d) I'm not sure

4. When are the Olympic Games officially closed?
   (a) After the IOC president makes his closing speech
   (b) After the Olympic flame is extinguished
   (c) After the medals are given for the final event
   (d) I'm not sure

Text 5

In 1989, the International Basketball Federation (FIBA) approved the rule allowing players in the professional National Basketball Association (NBA) teams to compete in international tournaments, including the Olympics. In the 1992 Olympics the US "Dream Team" won the gold medal with average winning margin of 44 points and without calling a time out. By this time, Soviet Union and Yugoslavia no longer existed, but their successors continued to be among the leading forces. Two newly-independent counties, Croatia and Lithuania, won the silver and bronze medals respectively.

The Americans repeated their victory in 1996 and 2000, but they were not as impressive as in 1992. The Olympics in 1996 was notable as Atlanta was the first Olympic host city to have its own NBA team, the Atlanta Hawks since the approval of competition by professional players. In the 1996 Olympics Yugoslavia won the silver medal with Lithuania winning the bronze. At the Sydney Olympic Games in 2000, France won the silver with Lithuania taking the bronze.

The dominance of the United States was interrupted in 2004 when the Americans suffered their third defeat in Olympic history to Argentina in the **semifinals**. The Argentineans went on to beat Italy in the final and become the fourth team to win the Olympic basketball title. However, the United States regained the gold medal in 2008.

The United States is by far the most successful nation in Olympic basketball. American men's teams have won 13 out of 16 tournaments they have participated in, including seven successive titles from 1936 to 1968. American women's teams have won six titles out of eight, including four in a row from 1996 to 2008.

Target word for scanning task: semifinals

Question for skimming and detailed reading tasks

1. When did the FIBA allow professional National Basketball Association (NBA) players to compete in international tournaments, including the Olympics?
   (a) 1989
   (b) 1992
   (c) 1996
   (d) I'm not sure

2. Which country won the bronze medal in Olympic basketball in the 2000 Games?
   (a) Yugoslavia
   (b) Lithuania
   (c) Soviet Union
   (d) I'm not sure

3. Which was the fourth country to win the gold medal in the Olympic basketball title?
   (a) Croatia
   (b) France
   (c) Argentina
   (d) I'm not sure

4. How many times did the American women's basketball team win in the Olympic Games?
   (a) 6 times
   (b) 8 times
   (c) 13 times
   (d) I'm not sure

Text 6

According to historical records, the first ancient Olympic Games can be traced back to 776 B.C. They were dedicated to the Olympian gods and were staged on the ancient plains of Olympia. They continued for nearly 12 centuries, until Emperor Theodosius decreed in 393 A.D. that all such "pagan cults" be banned.

Olympia, the site of the ancient Olympic Games, is in the western part of the Peloponnese which, according to Greek mythology, is the island of "Pelops", the founder of the Olympic Games. Imposing temples, votive buildings, elaborate shrines and ancient sporting facilities were combined in a site of unique natural and mystical beauty. Olympia functioned as a meeting place for worship and other religious and political practices as early as the 10th century B.C. The central part of Olympia was dominated by the majestic temple of Zeus, with the temple of Hera parallel to it.

The Olympic Games were closely linked to the religious festivals of the cult of Zeus, but were not an integral part of a rite. Indeed, they had a secular character and aimed to show the physical qualities and evolution of the performances accomplished by young people, as well as encouraging good relations between the cities of Greece. According to specialists, the Olympic Games owed their purity and importance to religion.

The Olympic victor received his first awards immediately after the competition. Following the announcement of the winner's name by the herald, a Hellanodikis (Greek judge) would place a palm branch in his hands, while the **spectators** cheered and threw flowers to him. Red ribbons were tied on his head and hands as a mark of victory.

Target word for scanning task: spectators

Question for skimming and detailed reading tasks

1. According to historical records, to when can the first ancient Olympic Games be traced back?
    (a) 512 BC
    (b) 776 BC
    (c) 393 AD
    (d) I'm not sure

2. Where is Olympia?
    (a) the northern part of the Peloponnese
    (b) the eastern part of the Peloponnese
    (c) the western part of the Peloponnese
    (d) I'm not sure

3. The religious festival of which god were the Olympic Games closely linked to, although not part of?
    (a) Zeus
    (b) Hera

(c) Athena
(d) I'm not sure

4. What mark of victory was tied on the head and hands of an Olympic victor?
   (a) An olive branch
   (b) Flowers
   (c) Red ribbons
   (d) I'm not sure

Appendix 12: Content, target words, and multiple choices questions for experiment 3 and experiment 4 (Thai version)

For practice trail

กีฬาโอลิมปิกเป็นงานแข่งขันกีฬาระดับนานาชาติครั้งสำคัญที่ครอบคลุมทั้งกีฬาฤดูร้อนและกีฬาฤดูหนาว มี นักกีฬาหลายพันคนเข้าร่วมในการแข่งขันกีฬาหลากหลายประเภท โอลิมปิกเกมส์ได้รับการยอมรับว่าเป็นงาน แข่งขันกีฬาครั้งสำคัญที่สุดของโลก โดยมีมากกว่า 200 ประเทศเข้าร่วม ปัจจุบันโอลิมปิกเกมส์จัดขึ้นทุก ๆ สี่ปีใน ฤดูกาลที่เหมาะสมสำหรับแต่ละกลุ่มกีฬา

โอลิมปิกเกมส์สมัยใหม่เริ่มต้นขึ้นในปี ค.ศ. 1896 เมื่อบารอน ปิแอร์ เดอ กูแบร์แตง ได้จัดตั้งคณะกรรมการ โอลิมปิกสากลขึ้น คณะกรรมการนี้ได้พัฒนาไปเป็นหน่วยงานปกครองหลักของกระบวนการโอลิมปิกโดยมีกฎ บัตรโอลิมปิก คอยวางนิยามโครงสร้างและแนวทางการดำเนินงาน กระบวนการโอลิมปิกประกอบด้วยสหพันธ์ กีฬานานาชาติของแต่ละชนิดกีฬา คณะกรรมการโอลิมปิกแห่งชาติของแต่ละประเทศ และคณะกรรมการจัดงาน ประจำการแข่งขันกีฬาโอลิมปิกครั้งนั้น ๆ ในฐานะที่คณะกรรมการโอลิมปิกสากลเป็นหน่วยงานปกครองหลัก คณะกรรมการฯ จึงมีหน้าที่เฟ้นหาเมืองที่เหมาะสมที่จะเป็นเจ้าภาพงานแข่งขันกีฬาโอลิมปิกแต่ละครั้ง จากนั้น เมืองเจ้าภาพจะมีหน้าที่จัดการแข่งขันกีฬาโอลิมปิกและหาทุนจัดพิธีเฉลิมฉลองตามที่กฎบัตรโอลิมปิกได้กำหนด ไว้

โปรแกรมการแข่งขันกีฬาโอลิมปิกจะประกอบด้วยกีฬาชนิดต่าง ๆ ที่จะใช้แข่งขันในกีฬาโอลิมปิก จะถูกกำหนด โดยคณะกรรมการโอลิมปิกสากล งานเฉลิมฉลองโอลิมปิกเกมส์นั้นครอบคลุมถึงพิธีการและสัญลักษณ์ต่าง ๆ อาทิเช่น ธงและคบเพลิงโอลิมปิก ตลอดจนพิธีเปิดการแข่งขัน และพิธีปิดการแข่งขัน มีนักกีฬามากกว่า 13,000 คนเข้าร่วมแข่งขันโอลิมปิกฤดูร้อนและฤดูหนาวใน 33 ชนิดกีฬา ในการแข่งขันเกือบ 400 รายการ ผู้เข้าแข่งขัน ที่ชนะเลิศอันดับหนึ่ง สองและสามในแต่ละรายการจะได้รับเหรียญโอลิมปิก ได้แก่เหรียญทอง เหรียญเงินและ เหรียญทองแดงตามลำดับ

วิวัฒนาการของกระบวนการโอลิมปิกในช่วงคริสต์ศตวรรษที่ 20 และ 21 ได้เปลี่ยนแปลงการแข่งขันกีฬา โอลิมปิกไปในหลาย ๆ ด้าน ตัวอย่างเช่น มีการริเริ่มกีฬาโอลิมปิกฤดูหนาว กีฬาโอลิมปิกสำหรับผู้พิการ หรือที่ เรียกกันว่าพาราลิมปิก และกีฬาโอลิมปิกสำหรับเยาวชน

กรุณาตอบคำถามด้วยการวงกลมรอบคำตอบที่ถูกต้อง

1. การแข่งขันกีฬาโอลิมปิกสมัยใหม่เกิดขึ้นครั้งแรกเมื่อใด
   (a) ค.ศ. 1876
   (b) ค.ศ. 1886
   (c) ค.ศ. 1896
   (d) ไม่แน่ใจ

2. ประเทศที่เข้าร่วมการแข่งขันกีฬาโอลิมปิกมีทั้งหมดประมาณกี่ประเทศ
   (a) มากกว่า 100 ประเทศ
   (b) มากกว่า 150 ประเทศ
   (c) มากกว่า 200 ประเทศ
   (d) ไม่แน่ใจ

3. บารอน ปิแอร์ เดอ กูแบร์แตง คือใคร
   (a) ผู้ก่อตั้งคณะกรรมการโอลิมปิกสากล
   (b) ผู้ก่อตั้งสหพันธ์กีฬานานาชาติ
   (c) ผู้ก่อตั้งคณะกรรมการโอลิมปิกแห่งชาติ
   (d) ไม่แน่ใจ

4. การแข่งขันกีฬาโอลิมปิกฤดูร้อนและโอลิมปิกฤดูหนาวมีการแข่งขันรวมกันทั้งหมดกี่รายการ
   (a) เกือบ 100 รายการ
   (b) เกือบ 400 รายการ
   (c) เกือบ 1000 รายการ
   (d) ไม่แน่ใจ

Text 1

กรีฑาประเภทวิ่งแข่ง เดินเร็ว กระโดดไกล และขว้างจักรเป็นกลุ่มชนิดกีฬาที่เก่าแก่ที่สุด โดยมีจุดเริ่มต้นมา
ตั้งแต่สมัยยุคก่อนประวัติศาสตร์ ปรากฏภาพของการแข่งขันกรีฑาอยู่บนหลุมศพชาวอียิปต์โบราณ ภาพการวิ่ง
แข่งและกระโดดสูงสามารถพบได้ตั้งแต่ 2,250 ปีก่อนคริสตกาล นอกจากนี้ยังมีเทลเชียนเกมส์ อันเป็นเทศกาล
โบราณของวัฒนธรรมเซลติกในไอร์แลนด์ที่พบว่าจัดขึ้นราว 1,800 ปีก่อนคริสตกาล การรวมตัวกันยาวนาน
สามสิบวันนั้น มีการวิ่งแข่งและขว้างหินรวมอยู่ในการแข่งขันด้วย

กีฬาในกรีฑาชนิดดั้งเดิมและชนิดเดียวในโอลิมปิกเมื่อ 776 ปีก่อนคริสตกาล ได้แก่ การวิ่งแข่งตามความยาว
ของสนามกีฬา หรือเรียกว่า สเตเดียน ต่อมาการแข่งขันชนิดนี้ขยายเป็นปัญจกีฬายุคโบราณ ที่เพิ่มขว้างจักร
และกระโดดสูงเข้ามา การแข่งขันกรีฑายังพบได้อีกในกีฬาระหว่างภูมิภาคของประเทศกรีซซึ่งจัดขึ้นหลังจากนั้น
เมื่อราวๆ 500 ปีก่อนคริสตกาล

คอทส์โวลด์ โอลิมปิกเกมส์ เทศกาลกีฬาที่จัดขึ้นในช่วงคริสต์ศตวรรษที่ 17 ที่ประเทศอังกฤษซึ่งรวมการแข่งขัน
กรีฑาประเภทการแข่งขว้างค้อน อีกทั้งยังมีการแข่งขันกีฬาประจำปีที่เรียกว่า โอลิมปิก เดอ ลา รีพับบิก ใน
ฝรั่งเศสช่วงยุคปฏิวัติจากปี ค.ศ. 1796 ถึง 1798 งานแข่งขันนี้เป็นรากฐานให้กับกีฬาโอลิมปิกฤดูร้อนสมัยใหม่
การแข่งขันที่เด่นที่สุดในงานได้แก่การวิ่งแข่ง นอกจากนี้ยังมีการรื้อฟื้นกีฬายุคโบราณของกรีซขึ้นอีกหลายชนิด
ปี ค.ศ. 1796 นับเป็นปีเริ่มต้นของการใช้ระบบเมตริกในการแข่งขันกีฬา

กรีฑาสมัยใหม่ถือกันว่าเริ่มต้นขึ้นที่วิทยาลัยทหารแซนด์เฮิร์ท ว่ากันว่าที่นี่เป็นที่แรกที่มีการจัดการแข่งขันกรีฑา
ในปี ค.ศ. 1812 และ ค.ศ. 1825 อย่างไรก็ตามไม่มีหลักฐานยืนยันที่ชัดเจน แต่ต่อจากนั้นไม่นานได้มีการบันทึก
ถึงมหกรรมการแข่งขันกรีฑาในร่มครั้งแรกของยุคสมัยใหม่ในช่วงทศวรรษที่ 1860 ซึ่งรวมถึงการแข่งขันที่แอ
ชเบิร์นแนม ฮอลล์ ที่กรุงลอนดอน

กรุณาตอบคำถามด้วยการวงกลมรอบคำตอบที่ถูกต้อง

1. หลักฐานที่บ่งบอกถึงงานแข่งขันกีฬากรีฑาที่เก่าแก่ที่สุดถูกค้นพบที่ใด
    (a) กรีซ
    (b) ฝรั่งเศส
    (c) อียิปต์
    (d) ไม่แน่ใจ

2. กีฬาชนิดใดเป็นกรีฑาชนิดดั้งเดิมและชนิดเดียวในการแข่งขันกีฬาโอลิมปิกครั้งแรก
    (a) วิ่งแข่งตามความยาวของสนามกีฬา
    (b) กระโดดสูง
    (c) ขว้างจักร
    (d) ไม่แน่ใจ

3. มีการนำระบบเมตริกมาใช้ในการแข่งขันกีฬาครั้งแรกเมื่อใด
    (a) 500 ปีก่อนคริสตกาลที่ประเทศกรีซ

(b) ค.ศ. 1796 ที่ประเทศฝรั่งเศส

(c) ค.ศ. 1812 ที่ประเทศอังกฤษ

(d) ไม่แน่ใจ

4. มหกรรมการแข่งขันกรีฑาในร่มครั้งแรกที่ได้รับการบันทึกไว้เกิดขึ้นเมื่อใด

(a) ในทศวรรษที่ 1700

(b) ในทศวรรษที่ 1790

(c) ในทศวรรษที่ 1860

(d) ไม่แน่ใจ

Text 2

สัญลักษณ์ของโอลิมปิกเกมส์ประกอบด้วยห่วง 5 สีที่คล้องกันอยู่ ได้แก่ ห่วงสีน้ำเงิน สีเหลือง สีดำ สีเขียวและสีแดง บนพื้นสีขาว สัญลักษณ์นี้ออกแบบขึ้นในปี ค.ศ. 1912 โดยบารอน ปิแอร์ เดอ กูแบร์แตง ผู้ก่อตั้งกีฬาโอลิมปิกเกมส์สมัยใหม่

การออกแบบสัญลักษณ์เช่นนี้ตั้งใจให้เป็นตัวแทนของห้าทวีปบนโลกที่มีมนุษย์อาศัยอยู่ ซึ่งได้มาประสานเป็นอันหนึ่งอันเดียวกันด้วยจิตวิญญาณโอลิมปิกและความมุ่งมั่นที่จะแข่งขันกันอย่างเต็มความสามารถ สีทั้งหกสีที่ถูกเลือก คือ สีน้ำเงิน สีเหลือง สีดำ สีเขียว สีแดงและพื้นหลังสีขาว เป็นชุดสีจากธงชาติของประเทศที่เข้าร่วมโอลิมปิกเกมส์ในช่วงที่สัญลักษณ์นี้เปิดตัว

ในบทความที่ตีพิมพ์ลงในนิตยสาร "โอลิมปิก รีวิว" ซึ่งเป็นนิตยสารทางการของคณะกรรมการโอลิมปิกสากล ที่ตีพิมพ์เมื่อเดือนพฤศจิกายน ค.ศ. 1992 นั้น โรเบิร์ต บาร์นี นักประวัติศาสตร์ชาวอเมริกัน ได้อธิบายว่าปิแอร์ เดอ กูแบร์แตงได้แนวคิดรูปห่วงคล้องกันมาจากตอนที่เขาเป็นหัวหน้าของสมาคมกีฬาฝรั่งเศสที่มีชื่อว่า USFSA สมาคมดังกล่าวก่อตั้งขึ้นด้วยการรวมสมาคมกีฬาสองสมาคมเข้าด้วยกันโดยมีสัญลักษณ์เป็นห่วงสองห่วงคล้องกัน ต่อมาหน่วยงานนี้ได้กลายมาเป็นตัวแทนของคณะกรรมการโอลิมปิกสากลในประเทศฝรั่งเศสจนถึงปี ค.ศ. 1925 เดิมทีรูปห่วงสองห่วงคล้องกันนั้นเป็นความคิดของจิตแพทย์ชาวสวิส คาร์ล ยุง สำหรับยุงแล้วห่วงเป็นตัวแทนของความต่อเนื่องและความเป็นมนุษย์

สัญลักษณ์ดังกล่าวได้ใช้ครั้งแรกเมื่อ ค.ศ. 1914 แต่หลังจากนั้นงดจัดโอลิมปิกเกมส์เนื่องด้วยสงครามโลกครั้งที่ 1 สัญลักษณ์นี้ได้รับการนำกลับมาใช้อีกครั้งอย่างเป็นทางการเมื่อมีการจัดโอลิมปิกเกมส์ขึ้นอีกภายหลังสงคราม โดยเปิดตัวครั้งแรกที่โอลิมปิกเกมส์ครั้งที่ 7 ปี ค.ศ. 1920 ที่เมืองแอนต์เวิร์ป ประเทศเบลเยียม ห่วงโอลิมปิกเริ่มโด่งดังและใช้กันอย่างแพร่หลายในช่วงก่อนถึงโอลิมปิกฤดูร้อน ปี ค.ศ. 1936 ที่กรุงเบอร์ลิน

กรุณาตอบคำถามด้วยการวงกลมรอบคำตอบที่ถูกต้อง

1. ใครคือผู้ออกแบบสัญลักษณ์ห่วงห้าสีบนพื้นขาวประจำการแข่งขันกีฬาโอลิมปิกสมัยใหม่
   - (a) บารอน ปิแอร์ เดอ กูแบร์แตง
   - (b) โรเบิร์ต บาร์นีย์
   - (c) คาร์ล ยุง
   - (d) ไม่แน่ใจ

2. นิตยสารทางการของคณะกรรมการโอลิมปิกสากลมีชื่อว่าอะไร
   - (a) IOC Magazine
   - (b) Olympic Magazine
   - (c) Olympic Revue
   - (d) ไม่แน่ใจ

3. สำหรับคาร์ล ยุง จิตแพทย์ชาวสวิส ความหมายของห่วงโอลิมปิกคืออะไร
   - (a) ความต่อเนื่องและความเป็นมนุษย์
   - (b) สปิริตโอลิมปิกและความเป็นมนุษย์
   - (c) ความต่อเนื่องและสปิริตโอลิมปิก
   - (d) ไม่แน่ใจ

4. ห่วงโอลิมปิกได้รับการเปิดตัวอย่างเป็นทางการเมื่อใด
   - (a) ค.ศ. 1912
   - (b) ค.ศ. 1914
   - (c) ค.ศ. 1920
   - (d) ไม่แน่ใจ

Text 3

ฟุตบอลไม่ได้เป็นส่วนหนึ่งของโอลิมปิกเกมส์สมัยใหม่ครั้งแรกในปี ค.ศ. 1896 เนื่องจากในขณะนั้นวงการ
ฟุตบอลระหว่างประเทศยังเพิ่งเริ่มต้น อย่างไรก็ดี บางข้อมูลกล่าวว่ามีการจัดการแข่งขันฟุตบอลอย่างไม่เป็น
ทางการขึ้นในโอลิมปิกเกมส์ครั้งแรก โดยทีมเอเธนส์ได้แพ้ให้กับทีมตัวแทนสเมอร์นา ซึ่งในขณะนั้นเป็นส่วนหนึ่ง
ของอาณาจักรออตโตมัน จากนั้นทีมสเมอร์นาได้พ่ายให้แก่ทีมจากเดนมาร์ก อย่างไรก็ดียังไม่มีความแน่ชัดว่า
การแข่งขันครั้งนี้เกิดขึ้นจริงหรือไม่

นักประวัติศาสตร์กีฬาโอลิมปิกกล่าวว่าไม่มีหลักฐานที่แน่ชัดจากปี ค.ศ. 1896 ที่บ่งบอกว่ามีการแข่งขันเกิดขึ้น
ระหว่างทีมกรีซและทีมเดนมาร์ก เป็นไปได้ว่าอาจมีความเข้าใจผิดเกิดขึ้นซึ่งได้ถูกเผยแพร่ต่อโดยไม่ได้รับการ
ตรวจสอบ

ทั้งนี้มีการแข่งขันฟุตบอลในโอลิมปิกเกมส์ปี ค.ศ. 1900 และ 1904 และในโอลิมปิกเกมส์ครั้งพิเศษในปี ค.ศ.
1906 แต่ทีมที่ลงแข่งขันยังเป็นทีมสโมสรหรือทีมชาติอย่างไม่เป็นทางการ ในปี ค.ศ. 1906 ทีมจากบริเตนใหญ่
เยอรมัน ออสเตรีย เนเธอร์แลนด์และฝรั่งเศสได้ถูกถอดออกจากการแข่งขันอย่างไม่เป็นทางการในครั้งนั้น ทำให้
เหลือแข่งขันกันเพียงทีมจากประเทศเดนมาร์ก สเมอร์นา เอเธนส์ และทีมเธสสาโลนิกิ มิวสิค คลับ โดยทีม
เดนมาร์กเอาชนะทีมเอเธนส์ไปได้ในรอบชิงชนะเลิศ ถึงแม้ว่าคณะกรรมการโอลิมปิกสากลจะถือว่าการแข่งขันใน
ปี ค.ศ. 1900 และ 1904 คือการแข่งขันอย่างเป็นทางการ แต่สหพันธ์ฟุตบอลระหว่างประเทศกลับไม่รับรองการ
แข่งขันทั้งสองครั้งนี้

หลังจากการนำกีฬาฟุตบอลเข้าสู่โอลิมปิกเกมส์ปี ค.ศ. 1908 ที่กรุงลอนดอน กีฬาฟุตบอลกลายเป็นส่วนสำคัญ
ประจำโอลิมปิกเกมส์ การแข่งขันทวีความสำคัญขึ้นตลอดช่วงทศวรรษที่ 1920 อย่างไรก็ดี ทศวรรษนี้ยังมีวันที่
ไม่ดีในประวัติศาสตร์ของกระบวนการโอลิมปิก เมื่อทีมเชคโกสโลวาเกียเดินออกจากสนามการแข่งขันในนัดชิง
เหรียญทองปี ค.ศ. 1920 เพราะไม่พอใจในตัวผู้ตัดสิน

กรุณาตอบคำถามด้วยการวงกลมรอบคำตอบที่ถูกต้อง

1. เพราะเหตุใดกีฬาฟุตบอลถึงไม่เป็นส่วนหนึ่งของโปรแกรมกีฬาอย่างเป็นทางการประจำโอลิมปิกเกมส์
   ปี ค.ศ. 1896
   (a) กีฬาฟุตบอลในระดับนานาชาติยังเพิ่งเริ่มต้น
   (b) ยังมีความสับสนว่าทีมตัวแทนประเทศหรือตัวแทนเมืองควรจะได้เข้าร่วมแข่งขัน
   (c) ในเวลานั้นยังไม่มีประเทศที่เล่นฟุตบอลมากเพียงพอ
   (d) ไม่แน่ใจ

2. สมาคมใดยอมรับการแข่งขันที่จัดขึ้นเมื่อปี ค.ศ. 1900 และ 1904 ว่าเป็นส่วนหนึ่งของโอลิมปิกเกมส์
   อย่างเป็นทางการ
   (a) คณะกรรมการโอลิมปิกสากล (IOC)
   (b) สหพันธ์ฟุตบอลระหว่างประเทศ (FIFA)
   (c) ทั้งคณะกรรมการโอลิมปิกสากล (IOC) และสหพันธ์ฟุตบอลระหว่างประเทศ (FIFA)
   (d) ไม่แน่ใจ

3. ในโอลิมปิกเกมส์ปี ค.ศ. 1906 ทีมเดนมาร์กเอาชนะทีมใดได้ในรอบชิงชนะเลิศ

    (a) บริเตนใหญ่

    (b) เอเธนส์

    (c) ฝรั่งเศส

    (d) ไม่แน่ใจ

4. ในปี ค.ศ. ใดที่ทีมเชโกสโลวาเกียเดินออกจากการแข่งขันฟุตบอลรอบชิงเหรียญทอง

    (a) ค.ศ. 1904

    (b) ค.ศ. 1908

    (c) ค.ศ. 1920

    (d) ไม่แน่ใจ

Text 4

พิธีเปิดกีฬาโอลิมปิกเกมส์จะต้องมีองค์ประกอบหลายประการตามที่กฎบัตรโอลิมปิกได้กำหนดไว้ พิธีการส่วน
ใหญ่ได้ถูกวางรากฐานไว้ในโอลิมปิกฤดูร้อน ปี ค.ศ. 1920 ที่เมืองแอนต์เวิร์ป โดยพิธีเปิดมักเริ่มต้นด้วยการเชิญ
ธงชาติและบรรเลงเพลงชาติของประเทศเจ้าภาพ จากนั้นประเทศเจ้าภาพจะจัดให้มีการแสดงทางดนตรี การร้อง
เพลง เต้นรำและการแสดงที่เป็นตัวแทนสื่อถึงวัฒนธรรมของชาตินั้นๆ

หลังส่วนของของศิลปะการแสดงจบลง คณะนักกีฬาจะเดินขบวนเข้าสู่สนามทีละประเทศ ตามธรรมเนียมแล้ว
กรีซจะเป็นประเทศแรกที่ก้าวเข้าสู่สนาม เพื่อเป็นการให้เกียรติแก่ถิ่นกำเนิดของกีฬาโอลิมปิก จากนั้นประเทศที่
เหลือจะเดินเข้าสู่สนามกีฬาตามลำดับตัวอักษรในภาษาของประเทศเจ้าภาพ โดยคณะนักกีฬาของประเทศ
เจ้าภาพจะเดินปิดท้ายขบวน เมื่อการกล่าวเปิดงานเสร็จสิ้นลง ถือว่าโอลิมปิกเกมส์ได้เริ่มต้นขึ้นอย่างเป็น
ทางการ สุดท้ายคบเพลิงโอลิมปิกจะได้รับการอัญเชิญเข้ามาในสนามกีฬาและส่งต่อไปจนกระทั่งถึงผู้วิ่งคบเพลิง
คนสุดท้าย ซึ่งมักจะเป็นนักกีฬาโอลิมปิกผู้มีชื่อเสียงและประสบความสำเร็จจากประเทศเจ้าภาพ ผู้วิ่งคบเพลิงจะ
จุดไฟโอลิมปิกในกระถางคบเพลิง

พิธีปิดการแข่งขันกีฬาโอลิมปิกเกมส์จะจัดขึ้นภายหลังการแข่งขันกีฬาทุกรายการเสร็จสิ้นแล้ว ผู้ถือธงจาก
ประเทศที่เข้าร่วมการแข่งขันแต่ละประเทศจะเดินเข้าสู่สนาม ตามด้วยนักกีฬาทุกคนที่จะเข้ามาพร้อมๆ กันโดย
ไม่มีการแบ่งแยกประเทศ มีการเชิญธงชาติสามธงขึ้นสู่ยอดเสาระหว่างการบรรเลงเพลงชาติของประเทศนั้นๆ
ได้แก่ธงของประเทศกรีซ เพื่อให้เกียรติประเทศต้นกำเนิดกีฬาโอลิมปิกเกมส์ ธงของประเทศเจ้าภาพในปีนั้น
และธงของประเทศที่จะเป็นเจ้าภาพกีฬาโอลิมปิกเกมส์ในครั้งต่อไป

ประธานคณะกรรมการจัดการแข่งขันและประธานคณะกรรมการโอลิมปิกสากล กล่าวปิดการแข่งขัน และไฟ
โอลิมปิกจะถูกดับลง นับเป็นการปิดฉากโอลิมปิกเกมส์อย่างเป็นทางการ

กรุณาตอบคำถามด้วยการวงกลมรอบคำตอบที่ถูกต้อง

1. พิธีการในพิธีเปิดการแข่งขันกีฬาโอลิมปิกถูกกำหนดขึ้นเมื่อใด
   (a) ค.ศ. 1896
   (b) ค.ศ. 1912
   (c) ค.ศ. 1920
   (d) ไม่แน่ใจ

2. การแข่งขันกีฬาโอลิมปิกถือว่าเริ่มต้นขึ้นอย่างเป็นทางการเมื่อใด
   (a) หลังคำกล่าวเปิดงานเสร็จสิ้น
   (b) หลังจุดคบเพลิงโอลิมปิกขึ้นที่กระถางคบเพลิงสนามแข่งขัน
   (c) หลังชักธงโอลิมปิกขึ้นสู่ยอดเสา
   (d) ไม่แน่ใจ

3. ธง 3 ผืนใดจะถูกเชิญขึ้นเสาระหว่างการบรรเลงเพลงชาติที่สอดคล้องกับธงนั้นๆ ในพิธีปิดการแข่งขัน
   (a) ธงชาติกรีซ ธงชาติของประเทศเจ้าภาพ และธงโอลิมปิก
   (b) ธงโอลิมปิก ธงชาติของประเทศเจ้าภาพ และธงชาติของประเทศเจ้าภาพครั้งถัดไป
   (c) ธงชาติกรีซ ธงชาติของประเทศเจ้าภาพ และธงชาติของประเทศเจ้าภาพครั้งถัดไป
   (d) ไม่แน่ใจ

4. การแข่งขันกีฬาโอลิมปิกจะปิดตัวลงอย่างเป็นทางการเมื่อใด
   (a) หลังประธานคณะกรรมการโอลิมปิกสากลกล่าวปิดงาน
   (b) หลังคบเพลิงโอลิมปิกดับลง
   (c) หลังการมอบเหรียญรางวัลของการแข่งขันสุดท้าย
   (d) ไม่แน่ใจ

Text 5

ในปี ค.ศ. 1989 สหพันธ์กีฬาบาสเกตบอลนานาชาติอนุมัติให้นักกีฬาอาชีพที่ลงแข่งในเอ็นบีเอสามารถลงแข่งใน
ระดับนานาชาติได้ ซึ่งรวมถึงโอลิมปิกเกมส์ด้วย ในการแข่งขันโอลิมปิกปี ค.ศ. 1992 "ดรีมทีม" จาก
สหรัฐอเมริกาสามารถคว้าเหรียญทองไปได้ด้วยการเอาชนะคู่แข่งด้วยแต้มห่างโดยเฉลี่ย 44 แต้มต่อนัด อีกทั้ง
ยังไม่มีการขอเวลานอก ในเวลานั้นไม่มีสหภาพโซเวียตหรือยูโกสลาเวียแล้ว แต่ทีมใหม่ที่มาแทนก็ยังคงมี
บทบาทสำคัญ โดยประเทศที่เพิ่งแยกตัวเป็นอิสระอย่างโครเอเชีย และลิทัวเนีย คว้าเหรียญเงินและเหรียญ
ทองแดงมาครองตามลำดับ

ชาวอเมริกันยังคงประสบความสำเร็จอย่างต่อเนื่องทั้งในปี ค.ศ. 1996 และ 2000 แต่ไม่มีครั้งใดที่จะขาดลอยเช่น
ในปี ค.ศ. 1992 โอลิมปิกเกมส์ปี ค.ศ. 1996 มีความน่าสนใจตรงที่เมืองแอตแลนต้าเป็นเมืองเจ้าภาพเมืองแรกที่
มีทีมเอ็นบีเอเป็นของตัวเองนับตั้งแต่มีการอนุญาตให้นักอาชีพสามารถลงแข่งได้ โดยทีมมีชื่อว่าแอตแลนตา
ฮอว์กส์ ในโอลิมปิกปี ค.ศ. 1996 ยูโกสลาเวียคว้าเหรียญเงิน ในขณะที่ลิทัวเนียได้เหรียญทองแดง ที่ซิดนีย์
โอลิมปิกเกมส์ในปี ค.ศ. 2000 ฝรั่งเศสได้เหรียญเงิน ส่วนลิทัวเนียได้เหรียญทองแดง

อเมริกาถูกโค่นแชมป์ในปี ค.ศ. 2004 เมื่อพ่ายให้กับอาร์เจนตินาในรอบรองชนะเลิศ ซึ่งนับเป็นความพ่ายแพ้ครั้ง
ที่สาม ตั้งแต่ลงแข่งในกีฬาโอลิมปิกมา โดยต่อมาทีมอาร์เจนตินาเอาชนะอิตาลีในรอบชิงชนะเลิศ และกลายเป็น
ทีมที่สี่ที่คว้าตำแหน่งแชมป์บาสเกตบอลโอลิมปิกไป อย่างไรก็ดี อเมริกาสามารถกลับมาทวงคืนเหรียญทองได้
ในปี ค.ศ. 2008

อเมริกาเป็นผู้นำที่โดดเด่นเป็นอย่างมากในกีฬาโอลิมปิกบาสเกตบอล ทีมชายจากอเมริกาได้คว้าตำแหน่งแชมป์
มาแล้ว 13 ครั้งจากการลงแข่งทั้งหมด 16 ครั้ง ซึ่งรวมถึงการได้แชมป์ติดต่อกันเจ็ดสมัยในช่วงปี ค.ศ. 1936 ถึง
1968 ในขณะที่ทีมหญิงจากอเมริกาคว้าแชมป์หกครั้งจากแปดครั้ง รวมถึงการเป็นแชมป์สี่สมัยติดต่อกันช่วงปี
ค.ศ. 1996 ถึง 2008

กรุณาตอบคำถามด้วยการวงกลมรอบคำตอบที่ถูกต้อง

1. สหพันธ์บาสเก็ตบอลนานาชาติ FIBA ยอมให้นักกีฬาจาก National Basketball Association (NBA)
    เข้าร่วมแข่งขันระหว่างประเทศรวมถึงกีฬาโอลิมปิกได้ตั้งแต่เมื่อใด

    (a) ค.ศ. 1989

    (b) ค.ศ. 1992

    (c) ค.ศ. 1996

    (d) ไม่แน่ใจ

2. ประเทศใดคว้ารางวัลเหรียญทองแดงจากการแข่งขันบาสเก็ตบอลโอลิมปิกในปี ค.ศ. 2000

    (a) ยูโกสลาเวีย

    (b) ลิทัวเนีย

    (c) สหภาพโซเวียต

    (d) ไม่แน่ใจ

3. ประเทศที่สี่ที่ได้เหรียญทองจากการแข่งขันกีฬาโอลิมปิกบาสเก็ตบอลคือ

(a) โครเอเชีย

(b) ฝรั่งเศส

(c) อาร์เจนตินา

(d) ไม่แน่ใจ

4. ทีมบาสเก็ตบอลหญิงของสหรัฐอเมริกาชนะการแข่งขันกีฬาโอลิมปิกรวมทั้งหมดกี่ครั้ง

(a) 6 ครั้ง

(b) 8 ครั้ง

(c) 13 ครั้ง

(d) ไม่แน่ใจ

Text 6

จากหลักฐานทางประวัติศาสตร์ โอลิมปิกเกมส์สมัยโบราณครั้งแรกที่พบมีขึ้นเมื่อ 776 ปีก่อนคริสตกาล งานแข่งขันนี้อุทิศให้แก่เทพเจ้ากรีกองค์ต่างๆ และจัดขึ้นบนที่ราบบนโอลิมเปีย การจัดการแข่งขันนี้จัดอย่างต่อเนื่องเป็นเวลาร่วม 12 ศตวรรษ จนกระทั่งจักรพรรดิธีโอโดเชียส ออกคำสั่งในปี ค.ศ. 393 ว่า "ลัทธินอกรีต" นี้ควรถูกยกเลิก

โอลิมเปีย อันเป็นสถานที่จัดการแข่งขันกีฬาโอลิมปิกยุคโบราณนั้นอยู่ทางฝั่งตะวันตกของเพโลพอนนีส ตำนานชาวกรีกว่าไว้ว่าบริเวณนี้เคยเป็นเกาะของเพลอบส์ ผู้เป็นผู้ริเริ่มกีฬาโอลิมปิก โอลิมเปียเป็นสถานที่ที่รวมเอาวิหารอันน่าเกรงขาม อาคารแก้บน ศาลบูชาเทพเจ้าอันละเอียดอ่อน และสถานที่จัดการแข่งขันกีฬาเข้าไว้ด้วยกันบนพื้นที่ที่อุดมด้วยความงดงามทางธรรมชาติอันดูลึกลับ โอลิมเปียเป็นทั้งจุดนัดพบสำหรับการบูชาเทพเจ้าและสำหรับกิจกรรมทางศาสนาและการเมืองอื่นๆ มาตั้งแต่สมัย 1,000 ปีก่อนคริสตกาล ศูนย์กลางของโอลิมเปียคือวิหารอันยิ่งใหญ่ของเทพซุส ซึ่งมีวิหารของเทพเฮราอยู่เคียงข้าง

โอลิมปิกเกมส์มีความสัมพันธ์อย่างใกล้ชิดกับเทศกาลทางศาสนาของลัทธิที่บูชาเทพซุส แต่ไม่ได้เป็นส่วนหนึ่งของกันและกัน อันที่จริงแล้วการแข่งขันกีฬานี้มีลักษณะเป็นกิจกรรมทางฆราวาสและเน้นแสดงออกให้เห็นถึงความสามารถทางกายและความก้าวหน้าทางด้านการแสดงของกลุ่มคนหนุ่มสาว อีกทั้งยังมีหน้าที่สานความสัมพันธ์ระหว่างเมืองต่างๆ ในประเทศกรีซ ผู้เชี่ยวชาญกล่าวว่าการแข่งขันกีฬาโอลิมปิกได้รับกลิ่นอายความบริสุทธิ์ศักดิ์สิทธิ์ต่อมาจากศาสนานี่เอง

ผู้ชนะโอลิมปิกจะได้รับรางวัลแรกของเขาทันทีที่จบการแข่งขัน เมื่อผู้ประกาศแจ้งชื่อผู้ชนะ เฮลลาโนดิไคส์ ซึ่งเป็นกลุ่มผู้ตัดสินของกรีซจะวางกิ่งต้นปาล์มลงบนมือของผู้ชนะ ระหว่างนั้นผู้ชมจะโห่ร้องและโยนดอกไม้ขึ้นมาให้ ริบบิ้นสีแดงจะถูกผูกไว้บนศีรษะและที่มือทั้งสองข้างเพื่อเป็นเครื่องหมายแทนชัยชนะ

กรุณาตอบคำถามด้วยการวงกลมรอบคำตอบที่ถูกต้อง

1. ตามบันทึกทางประวัติศาสตร์ โอลิมปิกเกมยุคโบราณครั้งแรกที่ค้นพบเกิดขึ้นเมื่อใด

    (a) 512 ปีก่อนคริสตกาล

    (b) 776 ปีก่อนคริสตกาล

    (c) ค.ศ. 393

    (d) ไม่แน่ใจ

2. โอลิมเปียอยู่ที่ใด

    (a) ตอนเหนือของคาบสมุทรเพโลพอนนีส

    (b) ฝั่งตะวันออกของคาบสมุทรเพโลพอนนีส

    (c) ฝั่งตะวันตกของคาบสมุทรเพโลพอนนีส

    (d) ไม่แน่ใจ

3. เทศกาลทางศาสนาของเทพเจ้าองค์ใดมีความเกี่ยวข้องอย่างใกล้ชิดกับโอลิมปิกเกมส์ถึงแม้จะไม่ได้เป็นส่วนหนึ่งของกันและกัน

    (a) ซุส

    (b) เฮรา

    (c) อธีนา

    (d) ไม่แน่ใจ

4. เครื่องหมายแทนชัยชนะชนิดใดจะถูกผูกไว้บนศีรษะและมือทั้งสองข้างของผู้ชนะโอลิมปิก

    (a) ช่อมะกอก

    (b) ดอกไม้

    (c) ริบบิ้นสีแดง

    (d) ไม่แน่ใจ

Appendix 13: Physical and visual fatigue questionnaire (English version)

Please indicate where on the scale between 'Strongly Agree' and 'Strongly Disagree' most reflects your visual fatigue.

1. After doing .... tasks on a recent text, my back and/or neck hurt from sitting in one position for so long.
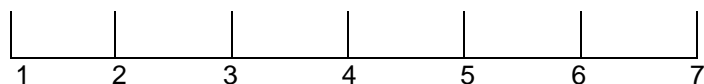
```
|    |    |    |    |    |    |
1    2    3    4    5    6    7
```
Strongly Disagree   Strongly Agree

2. Doing .... tasks on a recent text gives me a headache.
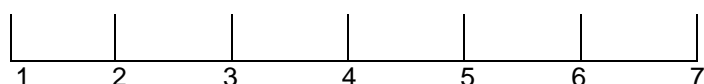
```
|    |    |    |    |    |    |
1    2    3    4    5    6    7
```
Strongly Disagree   Strongly Agree

3. After doing .... tasks on a recent text, my vision seems blurry when I look at distant objects.

```
|    |    |    |    |    |    |
1    2    3    4    5    6    7
```
Strongly Disagree   Strongly Agree

4. I feel mentally fatigued after doing .... tasks on a recent text.
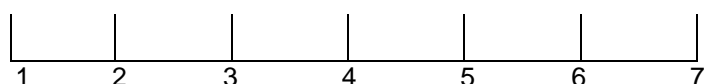
```
|    |    |    |    |    |    |
1    2    3    4    5    6    7
```
Strongly Disagree   Strongly Agree

5. After doing .... tasks on a recent text, my eyes feel strained.
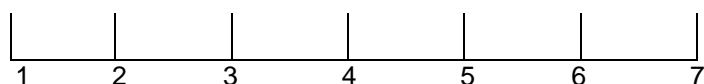
```
|    |    |    |    |    |    |
1    2    3    4    5    6    7
```
Strongly Disagree   Strongly Agree

6. Overall, doing .... tasks on a recent text makes me feel fatigued.

```
|    |    |    |    |    |    |
1    2    3    4    5    6    7
```
Strongly Disagree   Strongly Agree

Appendix 14: Physical and visual fatigue questionnaire (Thai version)

โปรดบ่งชี้ตำแหน่งบนมาตรประมาณค่าระหว่าง "เห็นด้วยเป็นอย่างยิ่ง" และ "ไม่เห็นด้วยเป็นอย่างยิ่ง" ที่สะท้อนความล้าทางตาของท่านมากที่สุด

1. หลังจากการ..........เนื้อหาล่าสุดที่พึ่งอ่านจบไป หลังและคอของฉันเจ็บ

```
     |       |       |       |       |       |       |
     1       2       3       4       5       6       7
```

ไม่เห็นด้วยเป็นอย่างยิ่ง   เห็นด้วยเป็นอย่างยิ่ง

2. . หลังจากการ..........เนื้อหาล่าสุดที่พึ่งอ่านจบไป ฉันรู้สึกปวดหัว

```
     |       |       |       |       |       |       |
     1       2       3       4       5       6       7
```

ไม่เห็นด้วยเป็นอย่างยิ่ง   เห็นด้วยเป็นอย่างยิ่ง

3. . หลังจากการ..........เนื้อหาล่าสุดที่พึ่งอ่านจบไป ตาของฉันเบลอเมื่อฉันมองวัตถุที่อยู่ระยะไกล

```
     |       |       |       |       |       |       |
     1       2       3       4       5       6       7
```
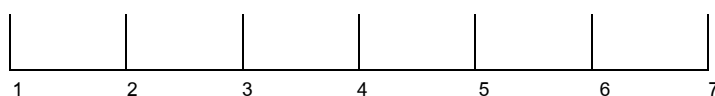
ไม่เห็นด้วยเป็นอย่างยิ่ง   เห็นด้วยเป็นอย่างยิ่ง

4. ฉันรู้สึกล้าทางสมอง หลังจากการ..........เนื้อหาล่าสุดที่พึ่งอ่านจบไป

```
     |       |       |       |       |       |       |
     1       2       3       4       5       6       7
```
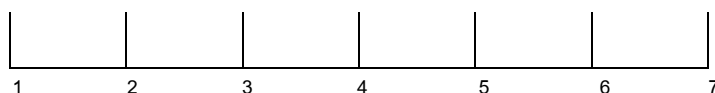
ไม่เห็นด้วยเป็นอย่างยิ่ง   เห็นด้วยเป็นอย่างยิ่ง

5. . หลังจากการ..........เนื้อหาล่าสุดที่พึ่งอ่านจบไป ตาของฉันมีความเครียด

```
     |       |       |       |       |       |       |
     1       2       3       4       5       6       7
```

ไม่เห็นด้วยเป็นอย่างยิ่ง   เห็นด้วยเป็นอย่างยิ่ง

6. .ในภาพรวม การ..........เนื้อหาล่าสุดที่พึ่งอ่านจบไปทำให้ฉันรู้สึกล้า

```
     |       |       |       |       |       |       |
     1       2       3       4       5       6       7
```

ไม่เห็นด้วยเป็นอย่างยิ่ง   เห็นด้วยเป็นอย่างยิ่ง

Appendix 15: Correlation between questions in Visual and Physical Fatigue for each Font Type and Font Size (N = 72)

Correlation/Font Type and Font Size	Serif 12 point	Serif 14 point	Serif 16 point	Sans serif 12point	Sans serif 14point	Sans serif 16point
Question 1/Question 6	.57**	.74**	.74**	.59**	.48**	.60**
Question 2/Question 6	.81**	.87**	.87**	.85**	.80**	.82**
Question 3/Question 6	.79**	.71**	.78**	.80**	.83**	.76**
Question 4/Question 6	.81**	.92**	.89**	.90**	.92**	.92**
Question 5/Question 6	.79**	.75**	.79**	.82**	.84**	.83**

* $p<.05$, ** $p<.01$

Appendix 16: Correlation in Ease of reading, Pleasantness of Reading, and Speed of Reading for each Font Type and Font Size

Correlation in Ease of reading, Pleasantness of Reading, and Speed of Reading for each Font Type (N = 72)

Correlation/Font Type	Serif	San serif
Ease-Pleasantness	.79**	.80**
Ease-Speed	.65**	.74**
Pleasantness-Speed	.66**	.72**

* p<.05, ** p<.01

Correlation in Ease of reading, Pleasantness of Reading, and Speed of Reading for each Font Size (N = 72)

Correlation/Font Size	12	14	16
Ease-Pleasantness	.54**	.56**	.52**
Ease-Speed	.53**	.47**	.51**
Pleasantness-Speed	.52**	.54**	.48**

* p<.05, ** p<.01

Appendix 17: Correlation between questions in Visual and Physical Fatigue for each Text Colour and Background Colour combination (N = 66)

Correlation/Colour	Black / White	White / black	Sepia / off-white
Question 1/Question 6	.64**	.62**	.65**
Question 2/Question 6	.76**	.86**	.82**
Question 3/Question 6	.78**	.76**	.79**
Question 4/Question 6	.83**	.88**	.89**
Question 5/Question 6	.86**	.84**	.80**

* $p<.05$, ** $p<.01$

Appendix 18: Correlation in Ease of reading, Pleasantness of Reading, and Speed of Reading for each Text Colour and Background Colour combination (N = 66)

Correlation/Colour	Black / White	White / black	Sepia / off-white
Ease-Pleasantness	.66**	.75**	.58**
Ease-Speed	.39**	.71**	.43**
Pleasantness-Speed	.53**	.63**	.59**

* p<.05, ** p<.01

Appendix 19: Correlation between questions in Visual and Physical Fatigue for each Reading task

Correlation between questions in Visual and Physical Fatigue for Searching for link word, Scanning, Skimming, and Detailed reading (N = 15) (Additional data collection)

Correlation/Reading task	Searching	Scanning	Skimming	Detailed reading
Question 1/Question 6	.75**	.54*	.70**	.48*
Question 2/Question 6	.78**	.71**	.67**	.54*
Question 3/Question 6	.80**	.62*	.87**	.61*
Question 4/Question 6	.79**	.81**	.87**	.92**
Question 5/Question 6	.90**	.75**	.87**	.76**

* p<.05, ** p<.01

Appendix 20: Participants' perceptions of their use of the reading types questionnaire

Please put the proportion of frequency which you use for scanning, skimming, and detailed readings on web pages

Scanning ……………………………..%

Skimming ……………………………%

Detailed reading …………………..%

Please put the proportion of time which you spent for scanning, skimming, and detailed reading on web pages

Scanning ……………………………..%

Skimming ……………………………%

Detailed reading …………………..%

Appendix 21: Methodologies for translation quality control

Translation quality control is important to cross-cultural research because the goal of cross-cultural translation is to achieve equivalence between two different languages (Lee et. al., 2009). Andriesen (2008) said that if all languages materials are not translated the exact same way and reflect the original, there is the risk that all answers according to these materials cannot be combined. In addition, data which collected from research materials with high quality of translation provide comparable data across countries and cultures (Mullis, Kelly, and Haley, 1996).

This section provides detail about quality control for preparation of the material in research.

1.1 Translation quality control methods

Mullis, Kelly, and Haley (1996) suggested that the translators and evaluators' work included the following:

- Identifying and minimizing cultural differences
- Finding equivalent words and phrases
- Making sure the reading level was the same in original as in target version
- Making sure the essential meaning of the items did not change
- Making sure the difficulty of the achievement items did not change
- Being aware of changes in layout due to translation.

Many researchers suggested techniques for control quality of translation (Brislin, 1970; Weeks, Swerissen, and Belfrage, 2007). The key translation techniques aim to maximize equivalence and minimize translation errors are:

- One-way translations
- Back translations
- Bilingual techniques
- Committee approach
- Pre-test procedures

1.1.1 One-way translation

Single bilingual person translates the sources language version into the target language. (de la Puente, Pan, and Rose, 2003). One – way translation technique is the most frequently used in public health cross-culture studies because this technique is simplicity, time efficiency, and low expense. However, one-way translated instrument is often result in low levels of validity and reliability (Week, Swerissen, and Belfrage, 2007).

1.1.2 Black translation

Translators who are target language native speakers translated content from the original language to target language. After that, the translated content is translated back into original language by independent translator or translators. Then the back translation is evaluated by comparing with original language (Andriesen, 2008; Weidmer, 1994). Back translation requires a few bilingual people but it requires high cost and amount of time (Andriesen, 2008).

1.1.3 Bilingual technique

Bilingual person completes a test in original language and target languages. Discrepancies in responses identify specific items which have not similarity. (Brislin, 1970; Smit et. al., 2006). Prince and Momnour (1967) explained that there were 3 methods in bilingual technique: (1) same bilingual participants completed first language materials then after one or two months in other languages. This method requires time (2) Bilingual participants were random into two groups to complete research materials; source and target language. The frequencies of response to the items could be compared. This method requires very large number of participants. (3) Bilingual participants were random to two groups. First group could be asked first half of questionnaires in source language and another could be asked the second half of questionnaires in target language. This method requires fewer participants than the second method. However, Prince and Momnour (1967) used 80 bilingual persons for their testing.

1.1.4 Committee approach

Committee approaches are used for translation and for translation assessment. For translation, group of bilingual people independently translates same materials from the source to the target language. The mistakes of one member can be caught by others on the committee. The next step, translators and a translation co-ordinator compare the independent translations, reconcile disparities and make a final version by select the best of the independent translations or follow other alternative comment in meeting (Brislin 1970; Harkness and Schoua-Glusberg, 1998; Smit et. al., 2006; Week, Swerissen, and Belfrage, 2007). For translation assessment, the committee members should proficiency in source and target languages and various related skills require for survey work (Guillemin et al., 1993).

1.2 Pre-test procedures

After a translation is completed, it is field-tested to ensure that future subjects will comprehend questions. Week, Swerissen, and Belfrage (2007) said that there are two methods. First is random-probe technique, interviewer selects a random item on questionnaires and as probing about that items (e.g., what do you mean?). When the researcher gets many responses, the

researcher knows about quality of the questionnaires. The second method is rating of questionnaires, interviewer ask samples to rate about how clear the question was to them.

However, some researchers used alternative technique in their research. Mullis, Kelly, and Haley (1996) suggested another technique which they used in translation of "Third International Mathematics and Science Study". In translation process, two Mathematics and two science specialists with excellent in English and target language independently translated the test items then there were two independent translations of each subject's test items. The next step, these two versions were compared by third person. If there were differences between two independent translations, the best version was selected. After that the translated content was sent to review by professional translator, first-language experience in target language, excellent knowledge of English, experience living and working in an English-language environment, and familiarity with the culture associated with the target language. The professional translator compared original version with translated version. If there were different, the professional translator gave a code for type of deviation and a code for severity of deviation and reported these in the translation verification report.

Type codes

The type codes, listed below, indicate what kind of change was made in the translation from the international version to the target language. Codes A through J refer to deviations in the text of an item; K through N refer to deviations in the graphics or layout of an item.

The type codes are:

A   Spelling

B   Grammar

C   Vocabulary

D   Incorrect number or value

E   Error in equation or numeric notation

F   Missing or additional text

G   Change in meaning

H   Change in level of reading difficulty

I   Tabs, alignment, or text layout

J   Other problem with the text

K   Labels are missing

L   Wrong picture or picture is missing

M   Picture has been modified

N   Labels have been modified.

Severity codes

The severity codes ranged from 1 (serious error) to 4 (acceptable adaptation).

1. Major Change or Error: This could affect the results and need to make corrections. Examples include an incorrect translation of text such that the answer is indicated by the question.

2. Minor Change or Error: This was to be corrected if possible, but would not affect the results. Examples include spelling errors that do not affect comprehension; misalignment of margins or tabs; incorrect font.

3. Suggestions for Alternative: The translation may have been adequate, but the verifier suggested a different wording for the item.

4. Acceptable Changes: The verifier identified acceptable changes and appropriate adaptations. An example is where a reference to winter was changed from January to July for the Southern Hemisphere.

After that the translation verification report was sent to review all major errors (severity code 1). and the process with progress until there is no major errors.

In conclusion, information in this section is used for control the quality of translation content and research materials in the studies in chapter 3, 5 - 6. As most of the studies in this thesis conducted with participants in two different countries and languages; Thailand and the UK. The reliable translation quality control ensure that data collected in the experiments were comparable between the two groups of participants.

For the current research, there were three procedures while three translators and three evaluators participated in translation quality control.

Procedures

Procedure 1: this procedure was used for controlling translation quality of material in the first group. There were 3 persons in the procedure; translator 1, translator 2, and evaluator 1.

Procedure 2: this procedure was used for checking translation quality of material in the second group. There were 3 persons in the procedure; translator 1, translator 3, and evaluator 2

Procedure 3: this procedure was used for evaluation of the equivalence between the English language web site and Thai language web site. This procedure was evaluated by evaluator 3.

Translators

Translator 1: a bilingual researcher who forward translated materials in both groups from English into Thai.

Translator 2: a bilingual Ph.D. student in TESOL/Applied Linguistics with first-language experience in Thai, familiarity with Thai culture, excellent knowledge of English and experience of translation and back translation who has lived and studied in UK for more than two years. Translator 2 translated the materials in the first group, Thai version, which were translated by Translator 1 back into English.

Translator 3: a bilingual Ph.D. student in the Department of English and Related Literature with first-language experience in Thai, familiarity with Thai culture, excellent knowledge of English, basic knowledge about Olympics and sports, and experience with translation who has lived and studied in the UK for more than two years. Translator 3 translated content in the second group from English into Thai.

Evaluators

Evaluator 1: a native English speaker, excellent in English and research, who has experienced living and working in the UK for at least 10 years. Evaluator 1 compared English language materials in the first group with the back translation version.

Evaluator 2: a bilingual Ph.D. student in the Department of Language and Linguistic Science with first-language experience in Thai, familiarity with Thai culture, excellent knowledge of English, basic knowledge about Olympics and sports, and experience with translation who has lived and studied in the UK for approximately two years. Evaluator 2 compared translated materials in the second group which were translated by translator 1 with another version which was translated by translator 2.

Evaluator 3: a bilingual person with first-language experience in Thai, familiarity with Thai culture, basic knowledge about creating web sites, and experience with the Internet for more than 10 years. Evaluator 3 compared the English language web site with the Thai language web site.

Procedure 1: Translation quality control procedure for materials in the first group

Translator 1 translated materials in the first group: informed-consent forms, pre-questionnaires, post-questionnaires, and tasks in the experiment from English into Thai language. The translated materials were provided to Translator 2, the back translator, for translation of the

materials in Thai back into English. Translator 2 was informed about the nature of the experiment and translation goal. The materials were listed below:

- Informed-consent form
- Pre-questionnaire
- Post-questionnaire
- Tasks in the experiment

Once back translation was completed the back translated materials and original English materials were passed to evaluator 1 for approval of the equivalence. The materials were listed below:

- Original informed-consent form
- Original pre-questionnaire
- Original post-questionnaire
- Original tasks in the experiment
- Back translated informed-consent form
- Back translated pre-questionnaire
- Back translated post-questionnaire
- Back translated tasks in the experiment
- Type codes and severity codes of Mullis, Kelly, and Haley (1996)

Evaluator 1 evaluated by comparing back translated materials with original English materials. When there were any different meanings between back translated and original English materials, Evaluator 1 assigned type codes and severity codes to each difference. The differences were checked by Translator 1. Translator 1 made the updated translation. Translator 2 was asked to back translate again and Evaluator 1 was asked to evaluate. The process ran until there was no difference. This meant that the materials in the first group in both languages were equivalent. The data collected by using both of these language materials can be pooled and the results compared.

Procedure 2: Translation quality control procedure for materials in the second group

The experiment web site contained a large amount of content and web pages which entailed that the translation quality control was separated into two steps. In the first step, translator 1 and translator 3 independently forwarded selected translated content on the experiment web site from English into Thai. The selected content was from 100 per cent of web pages which had the answers of the tasks and 50 per cent of other web pages. In the next step, if the similarity of translated content which was translated by translator 1 and translator 3 was more than 80 per cent, the remainder content was translated by translator 1. If the similarity was less than 80%, the remainder content required 2 forward translators.

Translators were informed about the nature of the experiment and translation goals before starting independent translation. The materials are listed below:

· Translator's works suggested by Mullis, Kelly, and Haley (1996)
· Selected content from web sites (100 per cent of web pages which had an answer and 50 per cent of related pages)

When translator 1 and translator 2 completed their independent translation, two versions of forward translation and related materials were passed to evaluator 2 for checking the similarity of the meaning of the content. The materials are listed below:

· Translator's works suggested by Mullis, Kelly, and Haley (1996)
· Translated content on web site (translated by translator 1)
· Translated content on web site (translated by translator 3)
· Translated tasks in the experiment and the answers

Evaluator 2 evaluated that the percentage of similarity of these two versions was 90% with no major differences between both versions, and therefore translation quality in this procedure was acceptable. Two independent translations were reconciled to be one version by agreement of translator 1 and translator 2 as a committee. The translator 1 translated the remainder content of the experiment web site.

Procedure 3: Web sites similarity control

The similarity of original web sites in English and the translated web site in Thai should be close to perfect. This was due to the need to ensure that data collected in the experiment were comparable between the UK and Thai participants.

After completion of the translation quality control in Procedure 2, the researcher used translated content for creating the translated web site. Both versions of the web site were checked for similarity by Evaluator 3 who was experienced with the Internet and had basic knowledge about creating web sites. Evaluator 3 checked the original web site and the translated web site for the similarity of each web page in terms of the following aspects:

- Presentation structure (e.g. top banner, top navigation bar, central content   area)
- Number of hypertext links in each web page
- Equivalence label of the links
- Colours on web site
- Number of menu items and menu item order
- Number of pictures and position
- Number of paragraphs of content in each page
- Text size and text layout

While checking similarity, Evaluator 3 also read all content on the web sites. Evaluator 3 reported the similarity and differences of each web page and spelling errors. The differences and spelling errors on the web page were immediately fixed by the researcher until Evaluator 3 reported no difference. Then Evaluator 3 checked the next web page. The process was repeated until the original web site and the translated web site were equivalent.

Appendix 22: Recommendations on presenting illustration and animation

Table 2.8: Recommendations on illustration and animation from the 8 sets of guidelines on web design for older people

Web design guidelines	Recommendation on line spacing
SPRY Foundation (1999)	The more realistic the illustration, the easier it is for a user to understand and follow. Animated illustrations provide an even greater means of retaining information. Any animated graphics, however, should always be accompanied by text or a text only option should be provided for those whose equipment cannot handle animations. Visual problems and older technology can also interfere with a user appreciation of animations on a web page. Slower modems, browsers and processors may not be able to access these files. The biggest problem, however, is that blinking text or repetitive motion can be so distracting to someone with visual or concentration issues that the pages content may not be located or understood.
Holt (2000)	For older adults, the more realistic the illustration, the clearer it will be to understand. Animated instructions provide an even greater means of retaining information; however, they should always be accompanied by text for those users unable to view graphic images. Animation is another fun piece of technology, but it can also make it harder for an older person to use your page. Both visual problems and older technology interfere with seeing many of the files. In addition, repetitive motion can be so distracting that users may miss your real content.
Zhao (2001)	Only necessary information should be presented on the screen and important information should be highlighted. Use only simple, highly relevant graphics. Animation, or any quickly flashing or blinking elements, are highly distracting to peripheral vision. They distract people's attention from focusing on the main information, as well as causing short-term memory loss, slower reading speed, and compromise reading comprehension. With the increased use of multiple advertising banners on Web pages, this can be a significant problem. One possible way to alleviate this problem is to allow users to pause or stop animation, flashing or blinking elements

AgeLight (2001)	Flashing or blinking graphics are highly distracting. For both new users and those with diminished peripheral vision, such as glaucoma or cataracts, such animation can be the difference between viewing a site and not. Excessive pop-up windows and ads banners have this same impact, distracting the reader and drawing attention to everything else.
NIA/NLM (2002)	Use text-relevant images only. Use short segments of animation, video, or audio to reduce download time on older computers.
AARP (2004)	Not mentioned
SilverWeb (2005, 2007)	Graphics should be relevant and not for decoration. No animation should be present. Avoid moving text
Webcredible (2006)	Not mentioned

References

Agelight. (2001). *Interface design guidelines for user of all age.* Retrieved October 23, 2012 from: www.agelight.com/webdocs/designguide.pdf

AARP Foundation. (2004). *Our work.* Retrieved July 15, 2014 from: http://www.aarp.org/aarp-foundation/our-work/

AARP Foundation. (2004). *Designing web sites for older adults : a review of recent research.* Retrieved June 14, 2011 from: http://www.aarp.org/olderwiserwired/oww/resources/a_review_of_recent_literature_2004.html.

Alzhemer's Society. (2012a). *Dementia 2012: a national challenge.* London: Alzhemer's Society.

Alzheimer's Society. (2012b). *Mild cognitive impairment.* Retrieved March 15, 2014 from: http://alzheimers.org.uk/site/scripts/download_info.php?fileID=1773

Andriesen, S. (2008). *Benefiting from back translations. Clinical trial management.* Retrieved September 24, 2012 from: http://www.medilingua.com/pdf/BackTranslationsICTSummer%202008.pdf

Arch, A. (2008). *Web accessibility for older users: A literature review W3C working draft.* Retrieved October 12, 2010 from: http://www.w3.org/TR/2008/WD-wai-age-literature-20080514/

Aula, A. (2005). User study on older adults' use of the Web and search engines. *Universal Access in the Information Society, 4*(1), 67-81.

Aula, A., and Käki, M. (2005). Less is more in Web search interfaces for older adults. *First Monday, 10*(7).

Bailey, Bob. (2004). *Segmenting Adult Web Users into Meaningful Age Categories. Presentation, Usability University July 2004; jointly sponsored by GSA, HHS and AARP.* Retrieved April 28, 2011, from http://assets.aarp.org/www.aarp.org_/articles/research/oww/university/Bailey_AgeCategories4.ppt

Barcelona Field Studies Centre. (2011). *Population dependency.* Retrieved October 30, 2012 from: http://geographyfieldwork.com/PopulationDependency.htm

Bernard, M., and Mills, M. (2000). So, what size and type of font should I use on my website. *Usability News, 2*(2), 1-5.

Bernard, M. L., Chaparro, B. S., Mills, M. M., and Halcomb, C. G. (2003). Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal of Human-Computer Studies, 59*(6), 823-835.

Bernard, M., Fernandez, M., and Hull, S. (2002). The Effects of Line Length on Children and Adults' Online Reading Performance. *Usability News.* July 2002, Vol. 4 Issue 2.

Bernard, M.L., Mills, M.M., Peterson, M., and Storrer, K. (2001). A comparison of popular online fonts: Which is best and when?. *Usability News, 3*(2).

Bernard, M., Liao, C. H., and Mills, M. (2001a). The effects of font type and size on the legibility and reading time of online text by older adults. In *CHI'01 Extended Abstracts on Human Factors in Computing Systems* (pp. 175-176). ACM.

Bernard, M., Liao, C. H., and Mills, M. (2001b). Determining the best online font for older adults. *Usability News, 3*(1), 2001.

Bernard, M., Lida, B., Riley, S., Hackler, T., and Janzen, K. (2002). A comparison of popular online fonts: Which size and type is best. *Usability News, 4*(1), 2002.

Bernard, M. L., Chaparro, B. S., Mills, M. M., and Halcomb, C. G. (2003). Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal of Human-Computer Studies, 59*(6), 823-835.

Beymer, D., Russell, D., and Orton, P. (2008). An eye tracking study of how font size and type influence online reading. In *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction-Volume 2* (pp. 15-18). British Computer Society.

Blaschke, C. M., Freddolino, P. P., and Mullen, E. E. (2009). Ageing and Technology: A Review of the Research Literature. *British Journal of Social Work*, *39*(4), 641-656.

Bohan, M., and Scarlett, D. (2003). Can expanding targets make object selection easier for older adults. *Usability News, 5*(1)

Boyarski, D., Neuwirth, C., Forlizzi, J., and Regli, S. H. (1998). A study of fonts designed for screen display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 87-94). ACM.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology,* 1(3), 185-216.

Burn, A-M. O. (2003). *Evaluating technology for elders: towards a measure of attitudes.* Ph.D. thesis, Department of Psychology, University of Hertfordshire.

Caldwell, B., Cooper, M., Reid, L. G., and Vanderheiden, G. (2008). *Web content accessibility guidelines 2.0', Web Accessibility Initiative (WAI), World Wide Web Consortium (W3C).* Retrieved October 23, 2012 from: http://www.w3.org /TR/WCAG20

Chadwick-Dias, A., McNulty, M., and Tullis, T. (2003). Web usability and age: how design changes can improve performance. In *ACM SIGCAPH Computers and the Physically Handicapped* (No. 73-74, pp. 30-37). ACM.

Chadwick-Dias, A., Tedesco, D., and Tullis, T. (2004). Older adults and web usability: is web experience the same as web expertise?. In *CHI'04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1391-1394). ACM.

Chaparro, A., Bohan, M., Fernandez, J., Choi, S. D., and Kattel, B. (1999). The Impact of Age on Computer Input Device Use: Psychophysical and Physiological Measures. *International Journal of Industrial Ergonomics, 24*(5), 503-513.

Chan, A. H. S., and Lee, P. S. K. (2005). Effect of display factors on Chinese reading times, comprehension scores and preferences. *Behaviour & Information Technology, 24*(2), 81-91.

Chan, A. H. S., Tsang, S. N. H., and Ng A. W. Y. (2014). Effects of Line length, Line Spacing, and Line Number on Proofreading Performance and Scrolling of Chinese Text. *Human Factors, 56*(3), 521-534.

Chou, K. L., and Chi, I. (2002). Successful aging among the young-old, old-old, and oldest-old Chinese. *The International Journal of Aging and Human Development, 54*(1), 1-14.

Chisholm, W., Vanderheiden, G., and Jacobs, I. (1999). *Web Content Accessibility Guidelines 1.0, Web Accessibility Initiative (WAI), World Wide Web Consortium (W3C).* Retrieved October 23, 2012 from: http:// www.w3.org/TR/WCAG10.

College of Optometrists. (2013). Computer and other screens. Retrieved March 27, 2013 from: http://lookafteryoureyes.org/eye-care/computers-and-other-screens/

Coyne, K. P., and Nielsen, J. (2002). Web usability for senior citizens. *Design guidelines based on usability studies with people age, 65.*

Czaja, S. J., and Moen, P. (2004). Technology and employment. *Technology and Adaptive Aging,* 150-178.

Czaja, S. J., and Sharit, J. (1998). Ability-Performance Relationships as a Function of Age and Task Experience for a Data Entry Task. *Journal of Experimental Psychology, 4*(4), 332-351.

Czaja, S. J., Sharit, J., Ownby, R., Roth, D. L., and Nair, S. (2001). Examining Age Differences in Performance of a Complex Information Search and Retrieval Task. *Psychology and Aging, 16*(4), 564-579.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly,* 319-340.

Davis, F. D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies, 38*(3), 475-487.

de la Puente, M., Pan, Y., and Rose, D. (2003). *An overview of proposed Census Bureau Guidelines for the translation of data collection instruments and supporting materials. Paper presented at the meeting of the Federal Committee on Statistical Methodology.* Retrieved October 23, 2012 from: http://www.fcsm.gov/03papers/delaPuente_Final.pdf

Dickinson, A., Eisma, R., Gregor, P., Syme, A., and Milne, S. (2005). Strategies for teaching older people to use the World Wide Web. *Universal Access in the Information Society, 4*(1), 3-15.

Dillon, A., Kleinman, L., Choi, G. O., and Bias, R. (2006, April). Visual search and reading tasks using ClearType and regular displays: two experiments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 503-511). ACM.

Dillon, A., Richardson, J., and McKnight, C. (1990). The effects of display size and text splitting on reading lengthy text from screen. *Behaviour & Information Technology, 9*(3), 215-227.

Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics, 35*(10), 1297-1326.

Dillon, A. (2004). *Designing usable electronic text: Ergonomic aspects of human information usage.* CRC Press.

Disability Rights Commission. (2004). *The Web: Access and inclusion for disabled people. A formal Investigation conducted by the Disability Rights Commission.* London: The Stationery Office.

Dror, I. E., Katona, M., and Mungur, K. (1998). Age differences in decision making: To take a risk or not?. *Gerontology, 44*(2), 67-71.

Dyson, M., and Haselgrove, M. (2000). The effects of reading speed and reading patterns on the understanding of text read from screen. *Journal of Research in Reading, 23*(2), 210-223.

Eurobarometer. (2008). *Data protection in the European union citizens' perceptions analytical report.* Retrieved September 1, 2011 from: http://ec.europa.eu/public _opinion/flash/fl_225_ en.pdf

Ferreira, M., and Kowal, P. (2006). A Minimum Data Set on Ageing and Older Persons in Sub-Saharan Africa: Process and Outcome. *African Population Studies. 21*(1), 19–36.

Fidgeon, T. (2006). *Usability for older web users.* Retrieved May 20, 2012 from: http://www.webcredible.co.uk/user-friendly-resources/web-usability/older-users.shtml

Fukuda, R., and Bubb, H. (2003). Eye Tracking Study on Web-Use: Comparison Between Younger and Elderly Users in Case of Search Task With Electronic Timetable Service. *PsychNology Journal, 1*(3), 202-228.

Garfein, A. J., and Herzog, A. R. (1995). Robust aging among the young-old, old-old, and oldest-old. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 50*(2), S77-S87.

Gao, Q., Sato, H., Rau, P. L. P., and Asano, Y. (2007). Design effective navigation tools for older web users. In *Human-Computer Interaction. Interaction Design and Usability* (pp. 765-773). Springer Berlin Heidelberg.

Global AgeWatch. (2013). *Healthy life expectancy at birth*. Retrieved July 15, 2014, From: http://www.helpage.org/global-agewatch/data/healthy-life-expectancy-at-birth/

Goodman-Deane, J., Keith, S., and Whitney, G. (2009). HCI and the older population. *Universal Access in the Information Society, 8*(1), 1-3.

Gradisar, M., Humar, I., and Turk, T. (2007). The legibility of colored web page texts. In *Information Technology Interfaces, 2007. ITI 2007. 29th International Conference* (pp. 233-238). IEEE.

Greco, M., Stucchi, N., Zavagno, D., and Marino, B. (2008). On the portability of computer-generated presentations: The effect of text-background colour combinations on text legibility. *Human Factors, 50*(5), 821-833.

Gregor, P., Newell, A. F., and Zajicek, M. (2002, July). Designing for dynamic diversity: interfaces for older people. In *Proceedings of the Fifth International ACM Conference on Assistive Technologies* (pp. 151-156). ACM.

Groff, L., Liao, C., Chaparro, B., and Chaparro, A. (1999). Exploring how the elderly use the web. *Usability News, 1*(2), 1-2.

Guillemin, F., Bombardier, C., and Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology, 46*(12), 1417-1432.

Gujar, A. U., Harrison, B. L., and Fishkin, K. P. (1998). A comparative evaluation of display technologies for reading. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 42, No. 6, pp. 527-531). SAGE Publications.

Hall, R. H., and Hanna, P. (2004). The impact of web page text-background colour combinations on readability, retention, aesthetics and behavioural intention. *Behaviour & Information Technology, 23*(3), 183-195.

Hanson, V. L. (2001). Web access for elderly citizens. In *Proceedings of the 2001 EC/NSF Workshop on Universal Accessibility of Ubiquitous Computing: Providing for the Elderly* (pp. 14-18). ACM.

Hanson, V. L. (2010). Influencing technology adoption by older adults. *Interacting with Computers, 22*(6), 502-509.

Harkness, J. A., and Schoua-Glusberg, A. (1998). Questionnaires in translation. *ZUMA-Nachrichten Spezial, 3*(1), 87-127.

Hassenzahl, M. & Tractinsky, N. (2006). User Experience – a research agenda. Behavior & Information Technology, 25(2), 91-97.

Hawthorn, D. (2003). How universal is good design for older users?. In *ACM SIGCAPH Computers and the Physically Handicapped* (No. 73-74, pp. 38-45). ACM.

Hearing Link. (2011). *Facts about deafness and hearing loss.* Retrieved July 15, 2014 from:http://www.hearinglink.org/facts-about-deafness-and-hearing-loss

Heinssen Jr, R. K., Glass, C. R., and Knight, L. A. (1987). Assessing computer anxiety: Development and validation of the computer anxiety rating scale. *Computers in Human Behavior, 3*(1), 49-59.

Hill, A., and Scharff, L. V. (1997). Readability of websites with various foreground/background color combinations, font types and word styles. In *Proceedings of 11th National Conference in Undergraduate Research* (Vol. 2, pp. 742-746).

Hodes, R., and Lindberg, D. A. B. (2002). *Making your web site senior friendly: a checklist.* Retrieved May 20, 2012 from: http://www.nlm.nih.gov/pubs/checklist.pdf.

Holt, B. (2000). *Creating senior-friendly web sites.* Retrieved October 23, 2012 from: http://www.allagewell.com/documents/seniorfriendlywebsites.pdf.

Holt, B., and Komlos-Weimer. M. (1999). *Older adults and the world wide web.* Retrieved October 23, 2012 from: http://www.spry.org/pdf/website_creators _guide.pdf.

Huey, R. W., Buckley, D. S., and Lerner, N. D. (1996). Audible Performance of Smoke Alarm Sounds. *International Journal of Industrial Ergonomics, 18*(1), 61-69.

Humar, I., and Turk, T. (2008). The impact of color combinations on the legibility of a Web page text presented on CRT displays. *International Journal of Industrial Ergonomics, 38*(11), 885-899.

Loyd, B. H., and Gressard, C. (1984). Reliability and factorial validity of computer attitude scales. *Educational and Psychological Measurement, 44*(2), 501-505.

Ilyas, M. (2012). A study of web accessibility barriers for older adults, and heuristics evaluation of email websites based on web accessibility heuristics for older adults by AARP. *J Emerging Trends in Computing and Information Sciences, 3*(5), 806-813.

International Standards Organisation. (1998). ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) Part 11: Guidance on Usability. ISO.

Jacko, J. A., Scott, I. U., Sainfort, F., Moloney, K. P., Kongnakorn, T., Zorich, B. S., and Emery, V. K. (2003). Effects of multimodal feedback on the performance of older adults with normal and impaired vision. In *Universal Access Theoretical Perspectives, Practice, and Experience* (pp. 3-22). Springer Berlin Heidelberg.

Kantner, L., and Rosenbaum, S. (2003). Usable computers for the elderly: Applying coaching experiences. In *Professional Communication Conference, 2003. IPCC 2003. Proceedings. IEEE International* (pp. 10). IEEE.

Keates, S., Clarkson, P. J., and Robinson, P. (2004). Design for participation: providing access to e-information for older adults. *Universal Access in the Information Society, 3*(2), 149-163.

Keates, S., and Trewin, S. (2005). Effect of age and Parkinson's disease on cursor positioning using a mouse. In *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 68-75). ACM.

Kim, J., Pan, Y., and McGrath, B. (2004). Personalization in digital television: Adaptation of pre-customized ui design. In *Proceedings of the 2nd European Conference on Interactive Television: Enhancing the Experience* (pp. 169-171).

Kline, P. (2000). *Handbook of psychological testing.* Routledge.

Kolers, P. A., Duchnicky, R. L., and Ferguson, D. C. (1981). Eye movement measurement of readability of CRT displays. *Human Factors*, 23(5), 517-527.

Kowal, P. R., and Peachey, K. (2001). Information needs for research, policy and action on ageing and older adults: Indicators for the Minimum Data Set Project on ageing: A critical review in sub-Saharan Africa. Geneva. *World Health Organization.*

Koyani, S., Adhmadi, M., Changkit, M., Harley, K., and Bailey, R. (2002). Older Users and the Web. In *Human Factors and Ergonomics Society 46th Annual Meeting.*

Kruk, R. S., and Muter, P. (1984). Reading of continuous text on video screens. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 26*(3), 339-345.

Kurniawan, S., and Zaphiris, P. (2005). Research-derived web design guidelines for older people. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility* (pp. 129-135). ACM.

Kurniawan, S. H., King, A., Evans, D. G., and Blenkhorn, P. L. (2006). Personalising web page presentation for older people. *Interacting with Computers, 18*(3), 457-477.

Lin, D. M. (2003). Hypertext for the Aged: Effects of Text Topologies. *Computers in Human Behavior*, *19*(2), 201-209.

Lines, L., and Elliman, T. (2007). *Delivering Inclusive Access for Disabled or Elderly Members of the Community: System Analysis - Focus Groups.* European Union Project Deliverable D2.2A. Brunel University, Kingston Lane, Uxbridge, UB8 3PH, UK.

Ling, J., and Van Schaik, P. (2006). The influence of font type and line length on visual search and information retrieval in web pages. *International Journal of Human-Computer Studies*, *64*(5), 395-404.

Ling, J., and van Schaik, P. (2007). The influence of line spacing and text alignment on visual search of web pages. *Displays, 28*(2), 60-67.

Lee, C., Li, D., Arai, S., and Puntillo, K. (2009). Ensuring Cross-Cultural Equivalence in Translation of Research Consents and Clinical Documents: A systematic process for translating English to Chinese. *Journal of Transcutural Nursing.* Vol. 20 No. 1, 77 - 82.

Ling, J., and Van Schaik, P. (2002). The effect of text and background colour on visual search of web pages. *Displays, 23*(5), 223-230.

Ling, J., and Van Schaik, P. (2004). The effects of link format and screen location on visual search of web pages. *Ergonomics, 47*(8), 907-921.

Ling, J., and Van Schaik, P. (2006). The influence of font type and line length on visual search and information retrieval in web pages. *International Journal of Human-Computer Studies*, *64*(5), 395-404.

Ling, J., and van Schaik, P. (2007). The influence of line spacing and text alignment on visual search of web pages. *Displays 28*(2), 60-67.

Maguire, M., and Pearce, L. (2001). An investigation into attitudes to, and experience of, Internet shopping. In *Proceedings of HCI International '01: the 9th International Conference on Human-Computer Interaction.* New Orleans, USA, pp. 1088-1092

Malik S. A. (2011). *Older people, mobile technology and culture: An investigation of appropriate methods and personas in Malaysia and the UK.* Ph.D. thesis Department of Computer Science, University of York.

Marquié, J. C., Jourdan-Boddaert, L., and Huet, N. (2002). Do older adults underestimate their actual computer knowledge?. *Behaviour & Information Technology*, *21*(4), 273-280.

Mehta, P. (2014). *Addressing hearing loss.* Retrieved July 15, 2014 from: http://www.ageuk.org.uk/professional-resources-home/knowledge-hub-evidence-statistics/improving-later-life-series/the-fourth-generation/addressing-hearing-loss-/

Meyer, B., Sit, R. A., Spaulding, V. A., Mead, S. E., and Walker, N. (1997). Age group differences in World Wide Web navigation. In *CHI'97 Extended Abstracts on Human Factors in Computing Systems* (pp. 295-296). ACM.

Moffatt, K. A., and McGrenere, J. (2007). Slipping and Drifting: Using Older Users to Uncover Pen-Based Target Acquisition Difficulties. In *Proceeding: ASSETS '07: Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility.* New York, NY, USA: ACM.

Moffatt, K., and McGrenere, J. (2009). Exploring methods to improve pen-based menu selection for younger and older adults. *ACM Transactions on Accessible Computing (TACCESS), 2*(1), 3.

Moore, J., and Matthews, J. (2004). Blind Leading the Blind: Theorizing a Web for the Visually Impaired. In *Proceedings* of *Usability Professionals' Association 2004.*

Monk, A., 2009. *Why do the older old have problems with new domestic technologies?*. Retrieved October 15, 2010 from: http://www.iden.org.uk/londonevent.asp.

Morrell, R. W. (2005). http://www. nihseniorhealth. gov: the process of construction and revision in the development of a model web site for use by older adults. *Universal Access in the Information Society*, *4*(1), 24-38.

Morrell, R. W., Mayhorn, C. B., and Bennett, J. (2000). A survey of World Wide Web use in middle-aged and older adults. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *42*(2), 175-182.

Moscicki, E. K., Elkins, E. F., Baurn, H. M., and McNarnara, P. M. (1985). Hearing Loss in the Elderly: An Epidemiologic Study of the Framingham Heart Study Cohort. *Ear and Hearing*, *6*(4), 184.

Mullis, I.V.S., Kelly, D.L., and Haley, K. (1996). *Translation Verification Procedures* in M.O. Martin and I.V.S. Mullis Third International Mathematics and Science Study: Quality Assurance in Data Collection. Chestnut Hill, MA: Boston College.

Muter, P., and Maurutto, P. (1991). Reading and skimming from computer screens and books: the paperless office revisited?. *Behaviour & Information Technology, 10*(4), 257-266.

National Institute of Neurological Disorders and Stroke. (2014). *NINDS Parkinson's disease information page.* Retrieved July 15, 2014 from: http://www.ninds.nih. gov/disorders/parkinsons_disease/parkinsons_disease.htm

NIA/NLM. (2002). *Making your web site senior friendly: a checklist.* Retrieved May 20, 2012 from: http://www.nlm.nih.gov/pubs/checklist.pdf

Nichols, T. A., Rogers, W. A., Fisk, A. D., and West, L. D. (2001). How old are your participants? An investigation of age classifications as reported in human factors. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 45, No. 3, pp. 260-261). SAGE Publications.

Nickell, G. S., and Pinto, J. N. (1986). The computer attitude scale. *Computers in Human Behavior, 2*(4), 301-306.

Nielsen, J. (2002). *Usability for senior citizens.* Retrieved May 20, 2012 from: http://www.useit.com/alertbox/seniors.html

O'Hara, K., and Sellen, A. (1997). A comparison of reading paper and online documents. *CHI'97.* 22 – 27 March 97.

O'Hara, K. (1996). Towards a Typology of Reading Goals. *Rank Xerox Research Centre Technical Report EPC-1996,* 107.

Ojanpää, H., and Näsänen, R. (2003). Effects of luminance and colour contrast on the search of information on display devices. *Displays, 24*(4-5), 167-178.

Parkinson's UK. (2014). *What is Parkinson's?.* Retrieved July 15, 2014 from: http://www.parkinsons.org.uk/content/what-parkinsons

Prince, R., and Mombour, W. (1967). A technique for improving linguistic equivalence in cross-cultural surveys. *International Journal of Social Psychiatry, 13*(3), 229-237.

Pearson, R., and van Schaik, P. (2003). The effect of spatial layout of and link colour in web pages on performance in a visual search task and an interactive search task. *International Journal of Human-Computer Studies 59*(3), 327-353.

Petrie, H., & Kheir, O. (2007). The relationship between accessibility and usability of websites. In Proceedings of the Special Interest Group on Computer-Human Interaction 2007 Conference on Human Factors. pp. 397 - 406.

Petrie, H., Hamilton, F., and King, N. (2004). Tension, what tension?: Website accessibility and visual design. In *Proceedings of the 2004 International Cross-Disciplinary Workshop on Web Accessibility (W4A)* (pp. 13-18). ACM.

Pfeil, U., Arjan, R., and Zaphiris, P. (2009). Age differences in online social networking–A study of user profiles and the social capital divide among teenagers and older users in MySpace. *Computers in Human Behavior, 25*(3), 643-654.

Redish, J. G., and Chisnell, D. (2004). *Designing web sites for older adults : a review of recent research.* Retrieved June 14, 2011, from http://www.aarp.org /olderwiserwired/oww/resources/a_review_ of_recent_literature_2004.html.

RNIB. (2009). *The economic impact of partial sight and blindness in the UK adult population.* Retrieved July 15, 2014 from: http://www.rnib.org.uk/sites/default /files/FSUK_Report_0.pdf

RNIB. (2013). *Age-related macular degeneration (AMD).* Retrieved July 15, 2014 from: http://www.rnib.org.uk/eye-health-eye-conditions-z-eye-conditions/age-related-macular-degeneration-amd

RNIB. (2014a). *Evidence-based review Older people.* Retrieved July 15, 2014 from: http://www.rnib.org.uk/sites/default/files/RNIB_Evidence_based_review_older_people_0.pdf

RNIB. (2014b). *Key information and statistics.* Retrieved July 15, 2014 from: http://www.rnib.org.uk/knowledge-and-research-hub/key-information-and-statistics

RNID. (2011). *Facts and figures on deafness and tinnitus, action on hearing loss information.* Retrieved July 15, 2014 from: http://www.actiononhearingloss.org.uk/ ~/media/Files/Factsheets/Deaf%20awareness/pdf/Facts%20and%20figures%20on %20deafness%20and%20tinnitus%20July%202011.ashx

Ryu, M., Kim, S., and Lee, E. (2009). Understanding the Factors Affecting Online Elderly User's Participation in Video UCC Services. *Computers in Human Behavior, 25*(3), 619-632.

Salvi, S. M., Akhtar, S., and Currie, Z. (2006). Ageing changes in the eye. *Postgraduate Medical Journal, 82*(971), 581-587.

Sayago, S., Camacho, L., and Blat, J. (2009). Evaluation of techniques defined in WCAG 2.0 with older people. In *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A)* (pp. 79-82). ACM.

Schwartz, S. (2012). *Age-related hearing loss.* Retrieved July 15, 2014 from: http://www.nlm.nih.gov/medlineplus/ency/article/001045.htm

Seidler, R. D., Bernard, J. A., Burutolu, T. B., Fling, B. W., Gordon, M. T., Gwin, J. T., and Lipps, D. B. (2010). Motor control and aging: links to age-related brain structural, functional, and biochemical effects. *Neuroscience & Biobehavioral Reviews, 34*(5), 721-733.

Shaikh A.D., and Chaparro, B.S. (2005). The effect of line length and passage type on reading speed, comprehension, and user satisfaction. In *1st Annual Symposium Graduate Research And Scholarly Projects* (p. 71). Wichita State University.

Shaikh, A. D. (2005). The effects of line length on reading online news. *Usability News, 7*(2), 1-4.

Sheedy, J. E., Subbaram, M. V., Zimmerman, A. B., and Hayes, J. R. (2005). Text legibility and the letter superiority effect. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 47*(4), 797-815.

SilverWeb. (2005). Research-derived web design guidelines for older people. In *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 129-135). ACM.

SilverWeb. (2007). A systematic approach to the development of research-based web design guidelines for older people. *Universal Access in the Information Society, 6*(1), 59-75.

Smit, J., van den Berg, C.E., Bekker, L-G., Seedat, S., and Stein, D.J. (2006). Translation and Cross-cutural Adaption of a mental health battery in an African setting. *African Health Science. Vol. 6 No. 4.* Retrieved October 23, 2012 from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC 1832071/

Smith, M. W., Sharit, J., and Czaja, S. J. (1999). Aging, Motor Control, and the Performance of Computer Mouse Task. *Human Factors, 41*(3), 389-397.

Snyder, H. L., Decker, J. J., Lloyd, C. J., and Dye, C. (1990). Effect of image polarity on VDT task performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 34, No. 19, pp. 1447-1451). SAGE Publications.

Spirduso, W. W., Francis, K. L., and MacRae, P. G. (2005). *Physical Dimensions of Aging* (2nd ed.). Champaign, IL: Human Kinetics.

SPRY Foundation. (1999). *Older adults and the world wide web.* Retrieved October 23, 2012 from: http://www.spry.org/pdf/website_creators_guide.pdf

Struve, D., and Wandke, H. (2009). Video modeling for training older adults to use new technologies. *ACM Transactions on Accessible Computing (TACCESS), 2*(1), 4.

Tractinsky, N., Katz, A. S., and Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, *13*(2), 127-145.

Tullis, T. S., Boynton, J. L., and Hersh, H. (1995). Readability of fonts in the windows environment. In *Conference Companion on Human Factors in Computing Systems* (pp. 127-128). ACM.

Turns, J., and Wagner, T. S. (2004). Characterizing audience for informational web site design. *Technical Communication, 51*(1), 68-85.

United Nations. (2002).*World population ageing: 1950-2050.* Retrieved April 20, 2011, from: http://www.un.org/ esa/population/publications/worldageing19502050/

United Nations. (2009).*World population ageing 2009.* Retrieved April 20, 2011, from: http://www.un.org/esa/population/publications/WPA2009/WPA2009_WorkingPaper .pdf

United Nations. (2013a). *World population prospects: the 2012 revision.* Retrieved July 15, 2014, from: http://esa.un.org/wpp/Documentation/pdf/WPP2012_ HIGHLIGHTS.pdf

United Nations. (2013b). *World population ageing 2013.* Retrieved July 15, 2014 from: http://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldP opulationAgeing2013.pdf

UNFPA Thailand. (2006). *Population aging in thailand: prognosis and policy response.* Retrieved April 20, 2011, from: http://thailand.unfpa.org/documents /thai_ageing_englishversion.pdf

van Schaik, P., and Ling, J. (2003). The effects of screen ratio and order on information retrieval in web pages. *Displays, 24*(4), 187-195.

van Schaik, P., and Ling, J. (2006). The effects of graphical display and screen ratio on information retrieval in web pages. *Computers in Human Behavior, 22*(5), 870-884.

Vorvick, J. L. (2012). *Hearing loss.* Retrieved July 15, 2014 from: http://www.nlm.nih.gov/medlineplus/ency/article/003044.htm

Wang, L., Sato, H., Jin, L., Rau, P. L. P., and Asano, Y. (2007). Perception of movements and transformations in flash animations of older adults. In *Human-Computer Interaction. Interaction Design and Usability* (pp. 966-975). Springer Berlin Heidelberg.

Wang, L., Sato, H., Rau, P. L. P., Fujimura, K., Gao, Q., and Asano, Y. (2008). Chinese text spacing on mobile phones for senior citizens. *Educational Gerontology*, *35*(1), 77-90.

WebCredible. (2006). *Usability for older web users.* Retrieved May 20, 2012 from: http://www.webcredible.co.uk/user-friendly-resources/web-usability/older-users.shtml

Weeks, A., Swerissen, H., and Belfrage, J. (2007). Issues, challenges, and solutions in translating study instruments. *Evaluation Review, 31*(2), 153-165.

Weidmer, B. (1994). *Issues and guidelines for translation in cross-cultural research.* Retrieved September 24, 2012 from: http://www.amstat.org/sections/srms/Proceedings/papers/1994_215.pdf

Wilkinson, R. T., and Robinshaw, H. M. (1987). Proof-reading: VDU and paper text compared for speed, accuracy and fatigue. *Behaviour & Information Technology, 6*(2), 125-133.

Wolters, M., Georgila, K., Moore, J. D., and MacPherson, S. E. (2009). Being old doesn't mean acting old: How older users interact with spoken dialog systems. *ACM Transactions on Accessible Computing (TACCESS), 2*(1), 2.

World Health Organization. (2000). *Information Needs for Research, Policy and Action on Ageing and Older Adults in Africa.* Report of a Workshop on Creating a Minimum Data Set (MDS) for Research, Policy and Action on Ageing and the Aged in Africa. Harare, Zimbabwe

World Health Organization. (2006). WHO Definitions and Meta-data. Retrieved Jan 10, 2015 from: http://www.who.int/whosis/whostat2006DefinitionsAndMetadata.pdf

World Health Organization. (2012). *Interesting facts about ageing.* Retrieved July 15, 2014 from: http://www.who.int/ageing/about/facts/en/

Wright, P., and Belt, S. (2001). Combining language and animation in multimedia instructions can help older people. *Gerontechnology, 1*(1), 60-62.

Wright, P., Belt, S., and John, C. (2004). Fancy graphics can deter older users: a comparison of two interfaces for exploring healthy lifestyle options. In *People and Computers XVII - Designing for Society* (pp. 315-325). Springer London.

Zaphiris, P., Kurniawan, S., and Ghiawadwala, M. (2007). A systematic approach to the development of research-based web design guidelines for older people. *Universal Access in the Information Society*, *6*(1), 59-75.

Zhao, H. (2001). Universal usability web design guidelines for the elderly (age 65 and older). *Universal Usability in Practice*.