A Multilingual Automated Web Usability Evaluation Agent

Maysoon Abulkhair

Ph. D.

The University of Sheffield

2004

A Multilingual Automated Web Usability Evaluation Agent

Maysoon Abulkhair

A Thesis Submitted for the Degree of Doctor of Philosophy

The Department of Computer Science

The University of Sheffield

February 2004

Acknowledgment

All Praise to Allah the Merciful Lord

I would like to express my appreciation and gratitude to my supervisor Doctor Siobhán North for her support, and valuable advice which helped me achieve my goal. I am also grateful to both Dr. Nigel Bevan and Professor Mike Holcombe for serving as members of my committee.

I cannot thank my parent enough for their love and continuous support, may Allah bless them.

A very special thank to my husband for his tremendous and endless support and encouragement, may Allah reward him for all his efforts.

I also would like to thank my children Abdullah, Majda, and Maria for being so patient throughout my research period. I ask Allah to guide them and help them lead a very bright, happy and successful life.

I am also thankful to my mother in-law, my three sisters, my two brothers, all my family and in-laws for their continuous support, love and prayers.

I would also like to thank Sulaiman Alharbi, Danyah Ismail, Ameen and Abdullah Basnawi and Bassam Mashaat for their efforts in distributing the user satisfactory testing.

Finally, I am sure many other people should be acknowledged here as well; to my friends, colleagues and those who supported me and helped me in any form, thank you so much.

Abstract

The research underlying this thesis explored the development of a customised, intelligent and automated approach to web usability evaluation. An extensive survey of existing web usability evaluation tools was carried out to identify to weaknesses that could be investigated. As result three different issues are addressed in this thesis:

- Improving and testing usability guidelines particularly for languages other than English;
- Customising the evaluation;
- Developing an intelligent (capable of learning) evaluation technique.

This thesis presents a new methodology that uses agent technology, which can act and interact on behalf of its owner (the webmaster), to evaluate web pages. The evaluation involves two kinds of customisation, one which reflects the users' tastes and the other the aims of the webmaster.

In investigating customisation of web pages to reflect users' tastes the research considered applying this multilingual interface agent approach to the evaluation of multilingual pages in scripts other than the usual Latin. But no guidelines appear to exist for such scripts thus the first difficulty in assessing non-English web pages is the lack of any reliable guidelines. In order to explore multilingual evaluation the researcher first had established guidelines and chose to investigate Arabic. As result usability guidelines for Arabic were established via usability testing. The guidelines are an interesting result of the research in themselves.

This thesis presents a set of usability guidelines appropriate for evaluating Arabic web pages produced by testing 196 Arabic users. Also, it validates some of the current usability guidelines for Latin scripts. An interesting variation appeared between the presentations of the two dissimilar scripts, these variations affect font size, emphasized text presentation, the number of links in the web page and the meanings associated with colours.

The second form of customisation is represented in the ability to modify the usability evaluation to reflect the webmaster's preferences. This requires an intelligent approach involving learning. Three different kinds of learning were considered; fuzzy average learning, fuzzy learning and Q-learning. All are examined in this thesis in order to identify the most appropriate approach to apply. As the multilingual interface agent learns form its webmaster, Q-learning produced the most accurate evaluation.

This thesis represents a useful first step towards multilingual, intelligent, automated web usability evaluation using an agent technique. The automated web usability, multilingual interface agent developed can be customised to suit its users and improve its evaluation in response to the needs of its owner.

Table of Contents

List of Figures

List of Tables

List of Abbreviation

WWW	World Wide Web.
CRTs	Cathode Ray Tubes.
RGB	Red, Green and Blue.
CHI	Computer Human Interaction.
HCI	Human- Computer Interaction.
UCD	User- Centred Design.
HMI	Human Machine Interface.
MUSiC	Metrics for Usability Standards in Computing.
WET	Web Event-logging Tool.
WebRemUSINE	Web Remote USer INterface Evaluator.
WebSat	Web Static Analyzer Tool.
HTML	HyperText Markup Language

Chapter One

Introduction

Due to the World Wide Web (WWW) a revolution in computing has occurred; data networks have proliferated and are still growing rapidly. Now, almost everyone wants the opportunity to use the huge amount of information the Internet provides (Huhns and Singh, 1998). Much information is available, but it has been presented in a variety of ways and that can affect the web pages' tasks, users and environment.

In order to achieve a certain degree of success, web developers must plan their web sites in two steps: firstly they should identify their development needs and goals; secondly, they should create the site specification (Lynch and Horton, 2001; Lengel, 2002). This specification includes what the developers intend to do, and why. In addition they must identify the technology needed to achieve their goals, and estimate how long the process will take (Lynch and Horton, 2001; Lengel, 2002). It is important to clarify the statement, or the basic idea, behind the construction of a web page and to develop consistent web sites and pages (Nielsen and Tahir, 2002).

Part of the web developers' job is to identify the targeted users and to discover their preferences, their likes and dislikes in web page layout. This has a great effect on the web page's acceptance (Sklar, 2000; Lengel, 2002); thus it is important to construct web pages that match the users' profiles. A wide variety of usability guidelines exists and has been established by different authors (Rivlin et al. 1990; ISO, 1991; Preece et al. 1994; Shneiderman, 1998; Rigden, 1999; Nielsen, 2000; Sklar, 2000; Ivory, 2001; Lynch and Horton, 2001; Ivory and Hearst, 2002; Ivory and Hearst, 2002; Lengel, 2002; Nielsen and Tahir, 2002; Rohn et al. 2002; Graham, 2003), each one of them is concerned with how to achieve the users' satisfaction and to present a usable web design.

Web professionals and experts have developed methodologies to detect web design problems and provide alternative solutions. In doing so they must measure web pages usability. They suggest different methods for measuring the usability in order to simplify and improve web designs (Ivory, 2001). Measuring the usability with the variety of methods is useful in discovering different problems in the design, since each method has the ability to process only certain criteria. Additionally, some of these usability measures are appropriate to investigate design problems only at specific stages.

Despite the work on usability measures, usability continues to be a serious problem for web developers. In an effort to address this problem, different automated and non-automated methods have been used. Given the vast increase in web pages it is only now that it is worth considering automated methods. Automated usability evaluation tools are similar to standard evaluation techniques; different techniques can discover several different kinds of problem (Nielsen, 1993; Nielsen, 1994; Nielsen, 1995; Ivory, 2001). Examples of this will be considered later in Chapter 3.

In spite of the improvement in automated usability evaluation, the field is still under-explored (Ivory 2001). The weaknesses are customisation, intelligence, and that evaluation can only be as good as the guidelines it applies. These are the areas that will be addressed in this thesis. The issues this study is investigating are:

- Improving and testing usability guidelines.
- Customising the evaluation.
- Developing an intelligent (capable of learning) evaluation technique.

Ivory (2001) claims that subjective measures such as user satisfaction is unlikely to be predictable by conventional automated methods. A customised approach is needed to overcome the unpredictable changes in the usability evaluation that have not yet been addressed in automatic web usability evaluation tools. Two types of customisation will be scrutinised in this research: customisation to the user group and customisation in the light of the webmaster's need.

In examining customisation for users, the researcher considered multilingual pages because most users are not English and so it is an obvious area to explore. No guidelines appear to exist for other scripts. The first difficulty in assessing non-English web pages is the lack of any reliable guidelines (see Chapter 2). One obvious weakness of the existing usability guidelines is that they have been designed for, and evaluated on, Latin scripts. In order to explore multilingual evaluation the researcher first has established guidelines and chosen Arabic to explore because:

1. It is widely used. Arabic is one of the six official languages approved by the United Nations (Nations 1999).
2. It has a different writing direction with totally different script.
3. No guidelines exist, which in itself makes it interesting.
4. No cultural studies exist either.

The researcher has developed guidelines for Arabic and used them to illustrate customisation in automated evaluation for, even though a side effect, the Arabic guidelines are in themselves a useful product of the research.

Lack of intelligence is another weakness of the majority of the existing tools. The intelligence is used in learning the webmaster's preferences and is the second type of customisation investigated in the study. Different intelligent and learning algorithms such as neural network, fuzzy logic, genetic algorithms and reinforcement learning are discussed in Chapter 4 and can successfully process individual problems. In order to develop an impartial system and to organise the flow and the processing of the data, the agent technique is used. Creating an impartial system requires sensing the input, processing the problem, producing appropriate actions, and adjusting future actions. It can contain one or more intelligent, or learning algorithms depending on the agent's targeted goals. In the field of web pages evaluation this seems an obvious technique to achieve customisation from the webmaster's point view.

In this study, several tasks must be achieved prior to the customisation of the usability measure such as web metrics computation, reporting the web page usability evaluation to the webmaster, learning from the webmaster's feedback for future evaluation. Different learning and intelligent algorithms are used to perform the usability evaluation and its customisation. The intelligence implemented in this study

uses the agent technique to control the performance of the analyses, evaluation and customisation of the assessment of the web page.

Several reasons stand behind the selection of more than one learning approach. The most important reason is because this researcher has not come across any previous studies that applied the agent technique to evaluate and customise the usability evaluation to webmasters' preferences and so the area was open to experiment. Additionally, establishing three learning approaches gives the opportunity to compare and contrast the agent's learning in order to come up with a solid conclusion to identify the most appropriate approach to apply.

1.1 Outline of the Thesis

Chapter one contains the *Introduction*, which describes the problem statement and highlights the research goals and rationale. The literature review is divided into three chapters. Then two chapters introduce the research methodology, followed by the results chapters and finally the conclusion.

Chapter two discusses the *Web Design Guidelines*. The discussion of these guidelines focuses particularly on the web page aspects under scrutiny in the rest of the dissertation.

Chapter three sheds light on the *Usability Evaluation Methods* and presents several usability evaluation techniques.

Chapter four provides a background for *Agent Technology*, and its intelligent and learning algorithms.

Chapter five describes the first part of the research methodology. This chapter describes the usability testing steps in detail and contains the specification of the questions involved.

Chapter six, the '*Multilingual automated web usability evaluation agent's Methodology'*, explains the second part of the methodology. This part describes the evaluation of the fuzzy model, the three learning approaches and the utilisation of feedback.

Chapter seven, '*Linking the Web Usability Guidelines with the User Satisfaction'*, associates the current usability recommendations with the findings from user satisfaction to validate the existing guidelines.

Chapter eight is '*Assessing the Agent's Evaluation'* and introduces the output of the evaluation. This chapter presents the first part of the agent's findings which contain the analysis and the evaluation.

Chapter nine is '*Evaluating the Agent's Learning'*. It presents the output of the agent's three learning approaches and then relates the closest one to the feedback.

Chapter ten is the '*Conclusion'*, and gives a comprehensive summary of the output of the research accompanied by some suggestions for future work.

Chapter Two

Web Usability Guidelines

2.1 Introduction

A wide variety of usability guidelines exist, and has been established by different authors (Rivlin et al. 1990; ISO, 1991; Preece et al. 1994; Shneiderman, 1998; Rigden, 1999; Nielsen, 2000; Sklar, 2000; Ivory, 2001; Lynch and Horton, 2001; Ivory and Hearst, 2002; Ivory and Hearst, 2002; Lengel, 2002; Nielsen and Tahir, 2002; Rohn et al. 2002; Graham, 2003). Each one of them focuses on how to satisfy users by presenting a usable web design. These guidelines address a wide range of web page design issues starting from the browser's title to the detail of web page features (Ivory, 2001). However, there is little consistency or overlap among the abundant usability guidelines (Ratner et al. 1996; Ivory, 2001). Additionally, contradictions exist between guidelines, which might be because of their different ages, the changes in the technology or because they might be appropriate only for specific group of users.

This chapter presents a set of relevant guidelines based on current usability studies and experts' recommendations in addition to other common recommendations for legible web interfaces. The aim of reviewing the usability guidelines literature is to quantify them, to determine their importance, and to focus on their relationships. The usability testing presented in Chapter 5 was used to validate a subset of the guidelines, and Chapter 7 to demonstrate the improved guidelines.

Each section in this chapter begins with a review of the common usability recommendations. These recommendations are followed by a discussion which compares different quantitative measures of the guidelines.

2.2 Web Page Scanning

Because of the massive variety of data that users can view in the WWW, users are accustomed to quickly scan web pages instead of reading every part of them (Nielsen, 2000; Sklar, 2000; Nielsen and Tahir, 2002). The first, and most important, aspect the users perceive is any oddities in the design (Sklar, 2000; Lynch and Horton, 2001). In other words, the web developer should be aware of colour combinations, the balance among different aspects, alignment and scrolling (Lengel, 2002), since they affect the scanability. Observing and scanning the content of the web page is performed in sequences of steps illustrated in Figure 2.1.

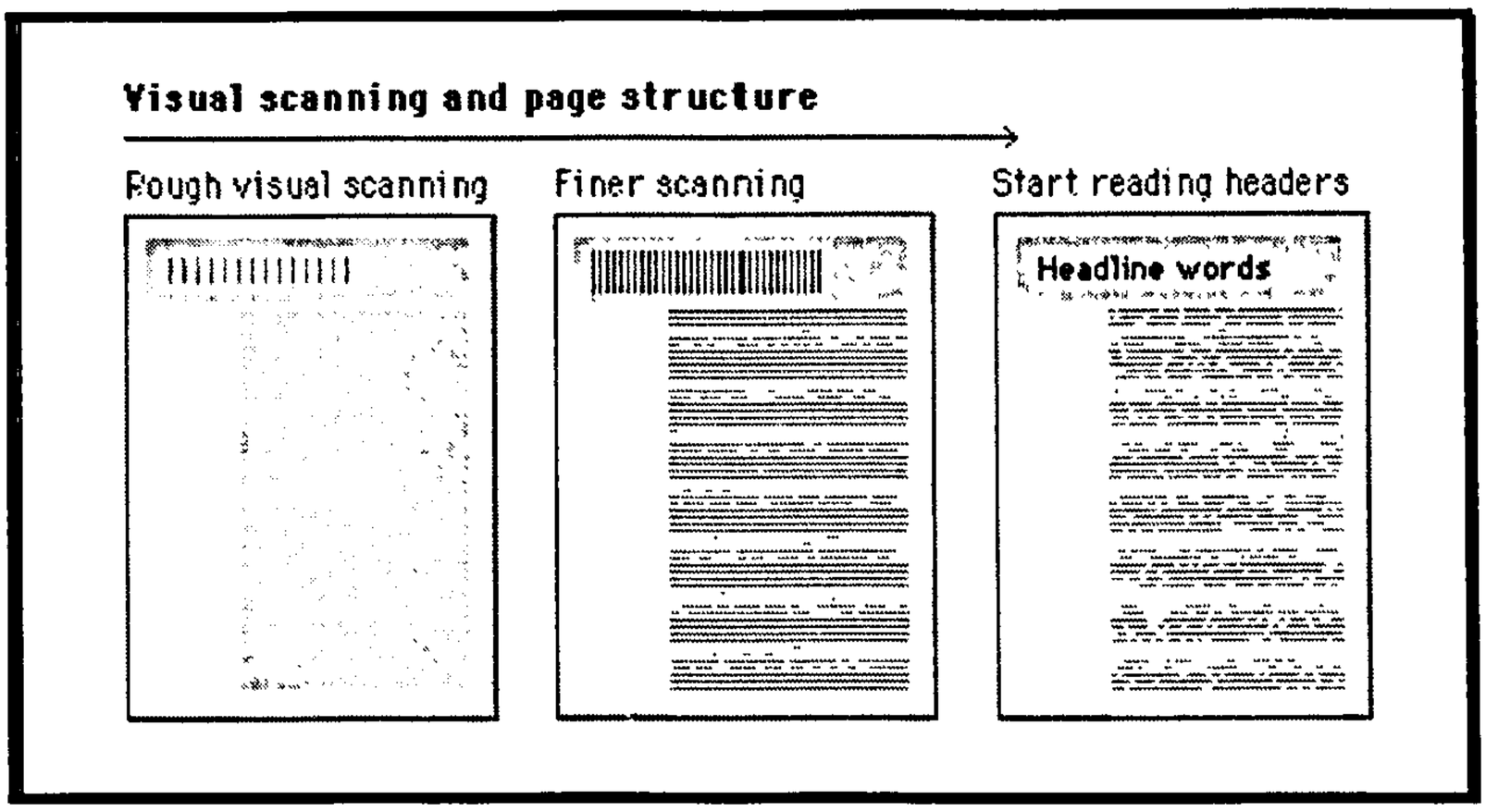


Figure 2.1: Users' Scanning Structure (Lynch and Horton, 2001)

Furthermore, web page consistency will affect readability. If the presentation of different web page components is not well organized, readability is reduced (Lynch and Horton, 2001). The basic concepts behind the construction and organisation of the data should be considered in order to produce a readable web page and to attract the users' eyes to certain data enclosed in the web page (Lengel, 2002).

2.3 Titles and Web Design Guidelines

The web page title is an important element in the web design as pointed out in (Ivory, 2001). Web pages differ from the usual printed pages because they are built to be searched; therfore their titles should be carefully selected to make them meaningful (Nielsen, 2000). The title is the first item that appears in the web browser window, it is considered as the main references to the web pages, and it can be used in the bookmark list and the history list (Nielsen, 2000; Nielsen and Tahir, 2002; Nielsen, 2003). On the web, users are always expected to scan rather than read (Nielsen, 2000; Nielsen and Tahir, 2002; Walton and Vukovic, 2003), consequently web titles should have limited length and be descriptive of the page's content (Nielsen and Tahir, 2002; Nielsen, 2003). Long web titles are useful for search engines but not for users; for this reason web usability experts do not recommend long titles. In order to successfully select the web page title experts offer some guidelines developers should follow.

- The ideal web page title length should be between 40 and 60 characters (Nielsen, 2000).
- Web page titles should not be more than 64 characters (Nielsen and Tahir, 2002).
- Web page titles should have two to six words (Nielsen, 2000; Ivory, 2001).
- Window titles should contain no more than eight words (Nielsen and Tahir, 2002).
- Web page titles should not contain words such as ".com, online, homepage and etc." (Nielsen and Tahir, 2002).

- Window titles should not start with words like "The" or "Welcome to" (Nielsen, 2002).

The first two guidelines are approximately compatible with each other; however, there is a difference between the third and the fourth guidelines. These differences were investigated by user satisfaction to produce impartial web title principles.

The above guidelines are intended to apply to the window or browser title, which can be extracted from the title tag in the HTML source code, and do not refer to the title of the document. The title of the web document or the main heading, which is located in the upper part of the web page, should also be carefully chosen as it is important to the reader but there are no specific guidelines that characterise web document titles in the literature. Research into this issue appears in Chapter 7.

2.4 Colour Combination and Web Usability Evaluation

Web users can view web pages through different screens with different facilities and which affect the appearance of colours. There are more than sixteen millions colours available to the vast majority of web users (Kerr, 2001), but many users can only view web pages on monitors with 256 colours (Lynch and Horton, 2001). These colours are supported by different graphics format like GIF, and also some of the major web browsers, for instance Internet Explorer and Netscape Navigator, effectively support them (Kerr, 2001; Antonacopoulos and Karatzas, 2002). They utilise colours as two different groups: the first one contains 40 different colours used by the operating system for icons and buttons. The other 216 colours can be used by the web designers and are called the safe-browser colours (Rigden, 1999; Sklar, 2000; Kerr, 2001).

Users' colour monitors are based on Cathode Ray Tubes (CRTs) which transmit light and use the Red, Green and Blue (RGB) additive colour model to display the colours used in the web design (Lynch and Horton, 2001; Antonacopoulos and Karatzas, 2002). In the RGB model the three basic colours can each have a value which varies between (0-255), and an individual colour is created by adding the values of the three different base colours.

Several colours can be chosen for each web design, and these colours can affect the usability of the web page, i.e. the selection colours can undermine or promote web design usability (Wilding, 1998).

The web designer should be careful with the text and background colour combination, as this too can affect the colour usability rate. Many web design experts recommend having a high contrast between the text and the background colour (Rigden, 1999; Nielsen, 2000; Nielsen and Tahir, 2002), or at least using sufficient contrast to ensure that the human eye can distinguish the difference between them (Sklar, 2000; Kerr, 2001; Lynch and Horton, 2001; Badre, 2002).

The colour contrast can help web developers to determine the details of the information perceived by the users (Mooney, 1998). Changes in the colour contrast flow between the text and background can reduce the users' concentration in extracting useful information from the less important data (Laurel, 1990; Head, 1999). The right selection of colour contrast can simplify the search process for the users

(Head, 1999) too. In order to determine efficient and effective colour contrast, several guidelines have been published which might help. These guidelines involve the number and type of colours within a web page and how they can influence the web page's usability.

Colour usability guidelines are a collection of different experts' viewpoints. They do not always agree. The conflicts are related to the age of the usability principles, the task served or cultural issues. The following are some of the colour usability principles, those that deal with the colour contrast between text and background colours.

- Get maximum contrast and the highest possible readability by selecting white background with black text (Shneiderman, 1998; Rigden, 1999; Nielsen, 2000; Ivory, 2001; Ivory and Hearst, 2001; Lynch and Horton, 2001; Lengel, 2002; Nielsen and Tahir, 2002) as many users with slight deficiency eye problems can at least recognise the difference between these two colours.
- Get less contrast and high readability by using any dark text colour over a white background (Nielsen and Tahir, 2002).
- The combination of a light text over a dark background has lower readability than the opposite. This combination selection is not recommended by many web usability experts (Rigden, 1999; Nielsen, 2000; Nielsen and Tahir, 2002), even though it was given the highest usability rate and recommended by (Rivlin et al. 1990; Preece et al. 1994).
- Select colours with sufficient contrast and which can be read by as many users as possible with different vision deficiencies (Sklar, 2000; Lynch and Horton, 2001), because similar intensities are harder to read (Dix et al. 1998).
- Avoid red and green, pink and green and blue and yellow colour combinations because they affect the readability of the web page in general and it might be impossible to read for colour blind users (Nielsen, 2000; Lynch and Horton, 2001).
- Avoid light text on a light background and dark text on dark background (Sklar, 2000).

It is clear that there is contradiction between number one and number three; therefore, this was tested by experiment.

After web designers have managed to select an efficient colour combination, tuning these colours within the web page is necessary to attract the users' eyes to the most important data in the web page. The colour alterations can add accent and interest to the display if, and only if, the number of colours and their combination are within reasonable bounds. The following are some of these recommendations.

- Minimize the number of colours used in the web page design (Ivory, 2001).
- Use conservative amounts of colour in each design because many colours can increase the search time (Preece et al. 1994).
- Limit the numbers of colours to four in a single display (Shneiderman, 1998).

- A good page measure is to have one to three colours within the web page (Ivory, 2001).
- Many colours can distract users from the content (Sklar, 2000).

It is obvious that there is difference in the range of colours that they recommend each web page should have. But all of them agree that too many colours could ruin the web design. Colour utilisation is not limited to aesthetics, adding interest for the users; instead they have different emotional meanings in different cultures (Mooney, 1998; Shneiderman, 1998).

Many of the previous colour guidelines are based on uncertain measurements, which makes them hard to evaluate. Also, because of the importance of the colours and the disadvantages of not using them properly, this research covers this web design aspect in more detail in the fuzzy colour model (Abulkhair and North, 2003) associated with user satisfaction.

After the immediate user recognition of the colours utilised in the web page, users scan the text. Usually, the scanning process starts from the upper part of the web page and moves downwards as explained in the next section.

2.5 Text Effects (Bold, Italic, Underlined)

Users can traverse web pages in several ways depending on their reading habits (Sklar, 2000). The track of users' eye movements can be classified either into a normal reading style starting from left to right and back again (see Figure 2.2) or as clockwise pattern (see Figure 2.3) (Sklar, 2000).

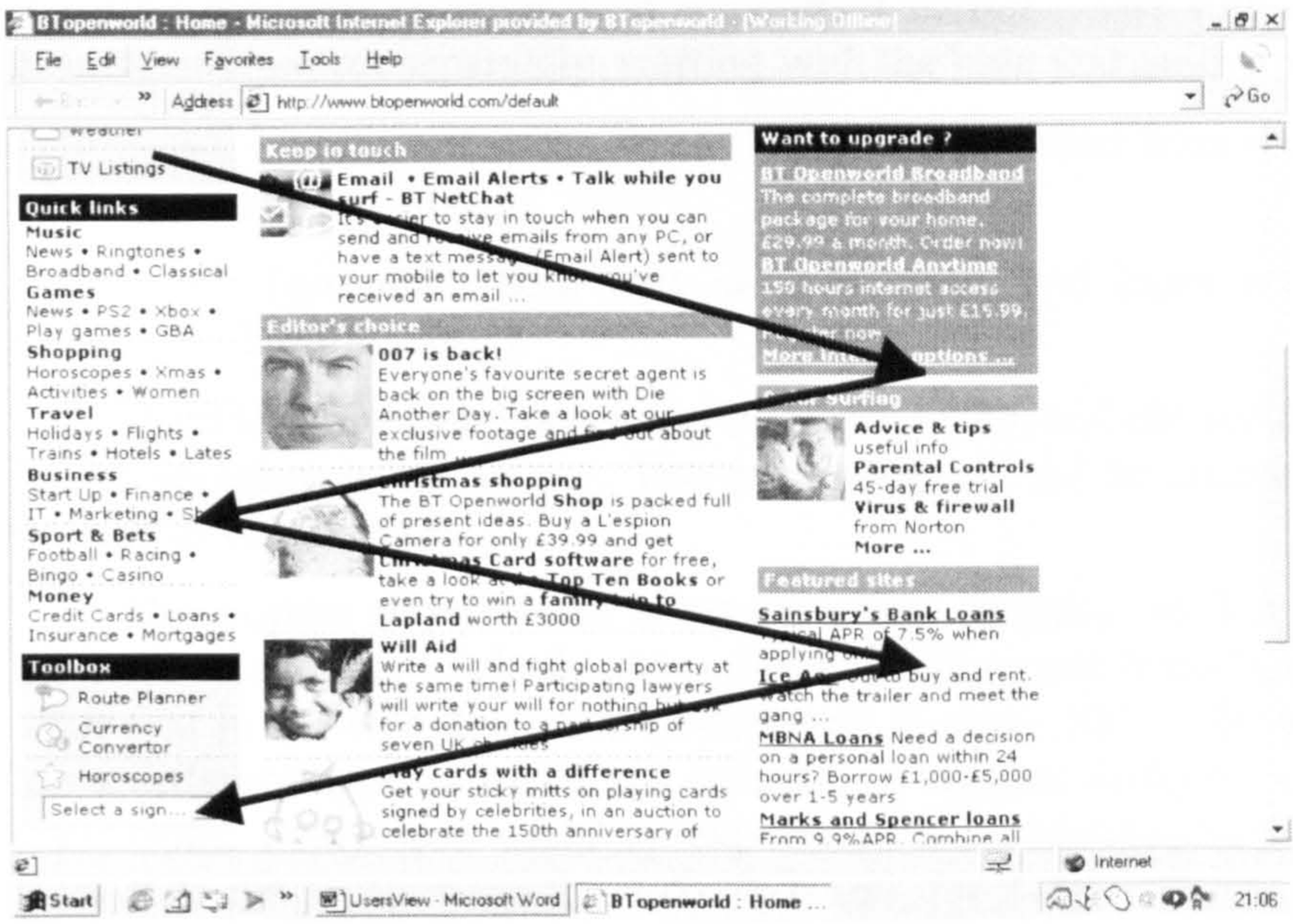


Figure 2.2: Paper based reading pattern for web viewing

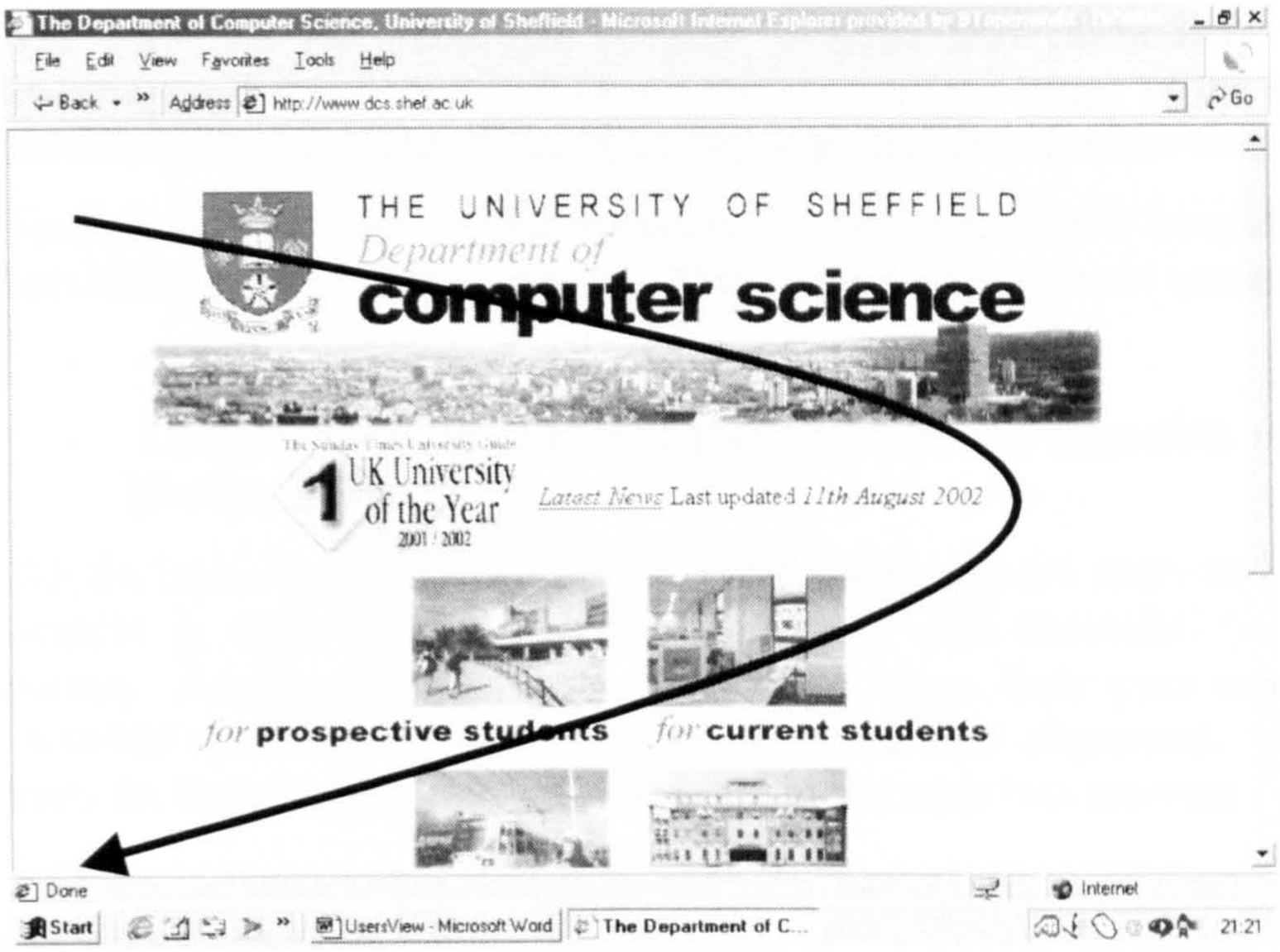


Figure 2.3: Clockwise pattern for web viewing

In both cases the users scan the web page from top to bottom searching for the desired data by concentrating on the emphasized text, which can be indicated by bold, italic, underlining, changed font size, face, colour or alignment. Each of these aspects has its own guidelines to improve its usage. Some of them are widely recommended whereas others are not. For instance, the use of bold text is recommended because it stands out from the rest of the text (Lynch and Horton, 2001). The guidelines for selecting the colours have been mentioned earlier. The following guidelines will discuss each aspect separately, starting with the bold text guidelines:

- Boldface text can be easily noticed and read from the screens (Lynch and Horton, 2001).
- Too much bold text lacks contrast and loses efficiency (Lynch and Horton, 2001).

Bold text gives an impression of a loud voice, and the redundant boldface usage is like shouting to the users. The developers should be careful when selecting the amount of boldface text.

Developers always want to attract users' attention so they might use italic text for emphasis because it looks attractive in printing and it contrasts in shape from the other text in the web document (Lynch and Horton, 2001). However, italic face text has a different appearance on the screen and because different screens have different resolutions, many web usability experts such as Nielsen, Sklar, Ivory, Lynch and Horton do not recommend the italic face text, for it also reduces web readability. The experts' recommendations for italic face usage can be summarized as follows:

- Avoid setting large blocks of text in italic (Lynch and Horton, 2001).
- Italic text is hard to read in paragraph format (Sklar, 2000).
- Avoid using italic text (Ivory, 2001).

The limit for the permitted amount of italic text contradicts with the third recommendation, because Ivory does not recommend the usage of the italic text. The determination of the amount of the italic text was decided after the user satisfaction.

Underlined text has special meaning in web documents because it represents hypertext links in many web browsers. Many web experts do not recommend using it.

- Avoid using underlined text (Ivory, 2001).
- Users can mix up the underline text with the hyperlink text (Lynch and Horton, 2001).

All the web experts' efforts aim to avoid confusing the users and to simplify the web content in order to increase the readability and accessibility of the desired information. As the users move to the reading stage, their eyes start to notice the more detailed specification including the font and text alignment. The underlying guidelines for these aspects will be discussed in the next two sections.

2.6 Font Specifications

Text plays an important role in making data available to the users, it involves two major issues: text content and text presentation. The presentation of text can be more important than the content, because even though web pages might have valuable data, poor font selection can deter users completely. There are some basic guidelines for producing efficient and effective fonts. These guidelines involve font colour, which was mentioned earlier, and face, size, capitalization, and alignment which will be discussed in the next sections.

2.6.1 Font Face Guidelines

Some font faces are hard to read on normal screens and some typefaces are more legible than others (Lynch and Horton, 2001). Also, some of these typefaces are legible in printed document but not on all screen resolutions. For example, Times New Roman is one of the most legible font faces on paper but on the screens its size looks small and its shape not as clear as it should be for maximum readability (Lynch and Horton, 2001). Font face guidelines can be summarised by the following:

- Use sans-serif font face for the small text size (Nielsen, 2000; Ivory, 2001; Lynch and Horton, 2001).
- Use serif typeface for the bigger text such as heading (Nielsen 2000; Ivory, 2001; Lynch and Horton, 2001).
- Use serif font for faster reading (Ivory, 2001).
- Use Time New Roman for long documents that need to be printed (Lynch and Horton, 2001).
- Georgia and Verdana offer excellent legibility for web pages (Ivory, 2001; Lynch and Horton, 2001).
- Use Verdana or Arial for headlines (Lynch and Horton, 2001).
- Use Georgia or Time New Roman for the body text (Lynch and Horton, 2001).

- The typeface choice should be based on the amount of web page content (Lynch and Horton, 2001).

The above guidelines favour the use of some font faces rather than others, because the character spacing in the recommended font is wide enough to be easily distinguished and read. The selected typeface should always be based on a font face designed for the computer screens (Ivory, 2001).

The number of the typefaces selected should not exceed two as pointed by (Ivory, 2001; Lynch and Horton, 2001), one for the headings and the other for the text body. Also, usability experts recommend using one or two different typeface families (Ivory, 2001).

2.6.2 Font Size Guidelines

The size of the font is important to fast scanning. The appearance of the font size differs from one browser to another, each of them has its own default font size (Lynch and Horton, 2001). Also an associated style sheet can affect the look of the font size. In order to achieve an appropriate layout with a specific font size for users with average vision, usability experts recommend the following:

- Use font size 14 point (or higher) for headings and displaying text (Ivory, 2001; Lynch and Horton, 2001).
- Use font size 12 point for body text (Ivory, 2001; Ivory and Hearst, 2001; Lynch and Horton, 2001; Lengel, 2002).
- Font size 10 and 11 point for the body text (Ivory, 2001).

In the above guidelines the font size in the body text varies from 10 to 12 point, depending on the underlying text's purpose. For instance, the footnotes should be the smallest font size, and the body text should be larger, etc.

2.6.3 Capitalize Guidelines

Some web developers rely on capitalized text to draw the users' attention toward a specific point as readers can recognize text in two ways, either by scanning the letters, or by recognizing the shape of the whole word (Lynch and Horton, 2001; Badre, 2002). However, capitalizing the text reduces web page readability, since the text appears as one block and the readers cannot distinguish different characters easily. This causes some difficulty in scanning it (Nielsen, 2000; Lynch and Horton, 2001; Badre, 2002; Nielsen and Tahir, 2002). Figure 2.4 illustrates the difference between the two letters' cases.

Thus, the majority of web usability experts do not recommend use of capital text even for small paragraphs (Nielsen, 2000; Ivory, 2001; Ivory and Hearst, 2001; Lynch and Horton, 2001; Badre, 2002; Lengel, 2002; Nielsen and Tahir, 2002). Users read all capitals text slower by ten percent than mixed cases (Nielsen, 2000). The following guidelines relate to capital font utilisation.

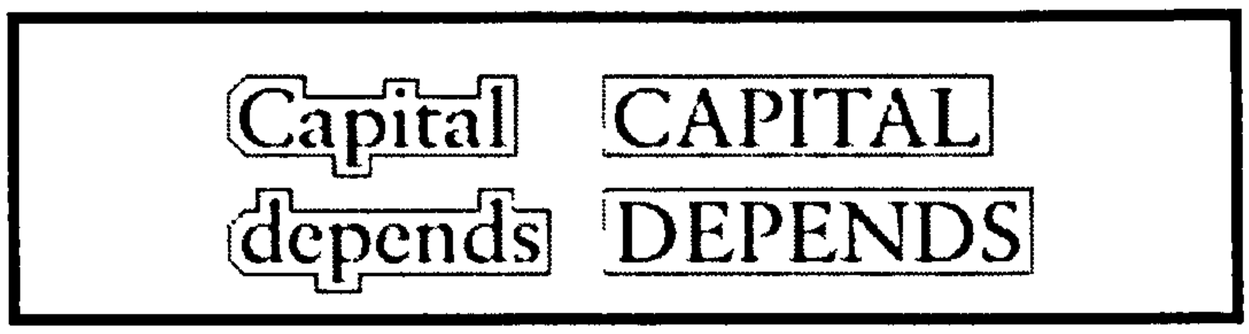


Figure 2.4: Differences of Letters' Cases (Lynch and Horton, 2001)

- Avoid using all caps text (Nielsen, 2000; Ivory, 2001; Ivory and Hearst, 2001; Lengel, 2002).
- The usage of text in capital case can be one or a few words (Badre, 2002).
- Titles, headings and subheadings can be presented using a few words in capital case (Badre, 2002).
- Avoid uppercase letters except for short headings (Lynch and Horton, 2001).
- Use uppercase letters sparingly or not at all (Nielsen and Tahir, 2002).

All the previous usability studies do not recommend the use of uppercase letters or restrict them to a few words if necessary. Capitalized text has a bad impact on readability and it makes the web pages look busy and loud (Nielsen and Tahir, 2002).

In reading capital text the users need to read letter by letter which is uncomfortable and more time consuming than mixed cases. All capitalized texts are considered as obstacles to fast reading.

All the font features should be carefully chosen to develop an efficient web page. Most of the guidelines do not recommend overuse of different fonts, but neither do they suggest any minimum number. By considering the minimum number of different features to be one, the researcher has run an experiment to determine the impact on users when using various numbers of fonts. The details and the result are discussed in Chapter 7.

2.7 Alignment Specification

As illustrated in Figures 2.2 and 2.3, users always scan from left to right (in English or other left to right languages (Sklar, 2000)); and if they start to read they look for the beginning of each line. The alignment of the text is very important in speeding up the users' reading and improving their ability to locate the desired piece of information. The following are some of the web usability guidelines for the text alignment.

- Web documents will suffer from using justified format (Lynch and Horton, 2001).
- Right and centred text alignments are both hard to read (Lynch and Horton, 2001).

- Centred and right text alignment should not be used for blocks of text (Nielsen, 2000; Ivory, 2001).
- Left justified is the most useful alignment and it is widely recommended (Ivory, 2001; Lynch and Horton, 2001).

The first guideline with respect to justification has several explanations. Firstly the text will appear as one block which can delay the reading process (Lynch and Horton, 2001). Secondly, justifying the text in web pages causes poor web page layout because of the word spacing problems it causes (Lynch and Horton, 2001). Users must search for the beginning of each line in the right and the centred alignment justification (Lynch and Horton, 2001).

A reasonable web heading justification that matches each of the above recommendations will be:

- Centred heading alignment for the justified text (Lynch and Horton, 2001).
- The best heading position is left alignment for the left justified text (Lynch and Horton, 2001).
- Centred and right heading alignment with the left justified text makes the display unbalanced (Lynch and Horton, 2001).
- Centred and right heading alignments are acceptable for a few lines (Nielsen, 2000).

All the previous guidelines indicate that right and centred text alignment can make the web document margins ragged, which makes them hard to read (Ivory, 2001; Lynch and Horton, 2001). Left justified is the most commonly recommended alignment for both the text and headings and also for making a consistent web document layout. But these guidelines will have to be tested to study their impact on the Arabic web documents.

2.8 Link specification

Links and their features differentiate web pages from the conventional printed documents. In the related literature there is no specific number of links each web page should have (Ivory, 2001). However, Nielsen recommends the minimization of links within a page (Nielsen, 2000) because the users might get confused with the outbound links. In other words, users might think that each link will take them out of the page they are scanning. The outbound links should be limited to the most relevant sites and not to all possible alternative sites in the web (Nielsen, 2000), since too many links of this kind can distract the users from the information provided in the web page.

There are many restrictions on the content of each link since, as pointed by (Ivory, 2001), the words contained in the linked text affect the information and navigation design. These restrictions involve the length of the text in a single link, the text content and the link colour. The following guidelines summarise these restrictions.

- Use two to four words in text link (Nielsen, 2000).

- A good page has one to three words in each link (Ivory, 2001).
- Links in a good page do not contain stop words (Ivory, 2001).
- Avoid using (click, click here, more and etc.) (Nielsen, 2000; Ivory, 2001; Lynch and Horton, 2001; Nielsen and Tahir, 2002).
- Use default browser colour for colouring the links (Nielsen, 2000; Ivory, 2001).
- Mark the visited links with a different colour to the unvisited (Nielsen and Tahir, 2002).

It is clear from the above link guidelines that all of them aim to produce meaningful descriptive links.

2.9 Line Length Specification

As users find the web content interesting and their eyes sweep over the page, they can take in most of the web page content (Sklar, 2000) if the line length is reasonable. That means the line length should be short enough to retain the users' eyes, because long lines might make them lose concentration in scanning. The ideal line length for text layout is based on the physiological features of human eyes (Lynch and Horton, 2001). There is a certain vision arc, limited to few inches, that users can easily scan without either moving their neck or overworking their eyes muscles (Lynch and Horton, 2001). There are some usability guidelines that help in structuring web page lines:

- The line length should not be more than 35 characters for web pages intended to be scanned (Badre, 2002).
- Use up to 75 characters for web pages that need to be read word by word (Badre, 2002).
- Use about ten to twelve words per line (Lynch and Horton, 2001).
- The best line length for readability is 10 to 12 words per line (Lengel, 2002).

A moderate line length can significantly increase the legibility of the web page, because if the users' eyes traverse large distance on a page, they are likely to lose the continuity in scanning (Lynch and Horton, 2001). Also, users need to hunt for the beginning of the next line each time they change line (Lynch and Horton, 2001) and lines that are bigger than the actual screen width require users to scroll horizontally, and produce an inconvenient display.

Horizontal scrolling is one of the most important features that web developers should avoid in their design. Many authors agree that horizontal scrolling has a negative impact on the usability rate (Lynch and Horton, 2001) and has been described as annoying feature (Sklar, 2000). It increases the users' memory load (Nielsen, 2000; Lynch and Horton, 2001) and users will miss seeing the content that scrolled off the screen (Lynch and Horton, 2001; Nielsen and Tahir, 2002).

2.10 Page length Specification

Researchers have noticed that vertical and horizontal scrolling both reduce web usability for reasons mentioned earlier. Additionally, with vertical scrolling users can lose the structure of the document because they might see different screens with different data content (Lynch and Horton, 2001) as they scroll downwards. However, these difficulties can be reduced, if the developers manage to organise their screen's display. There are some guidelines for managing the length of the web pages.

- Create a friendly web page by adding "jump to top of the page" after each section end (Lynch and Horton, 2001).
- Divide the web page into sub-pages that each carry certain topic (Lynch and Horton, 2001).
- Give the users a chance to navigate the content by providing a menu (Nielsen, 2000; Lynch and Horton, 2001).

The latest is a useful guideline, which can build a strong visual structure and add contrast to the display (Lynch and Horton, 2001). Using menus to organise the web pages make the information easier to locate and increases the page's legibility (Lynch and Horton, 2001). Figure 2.5 illustrates one of the menu organising methodologies.

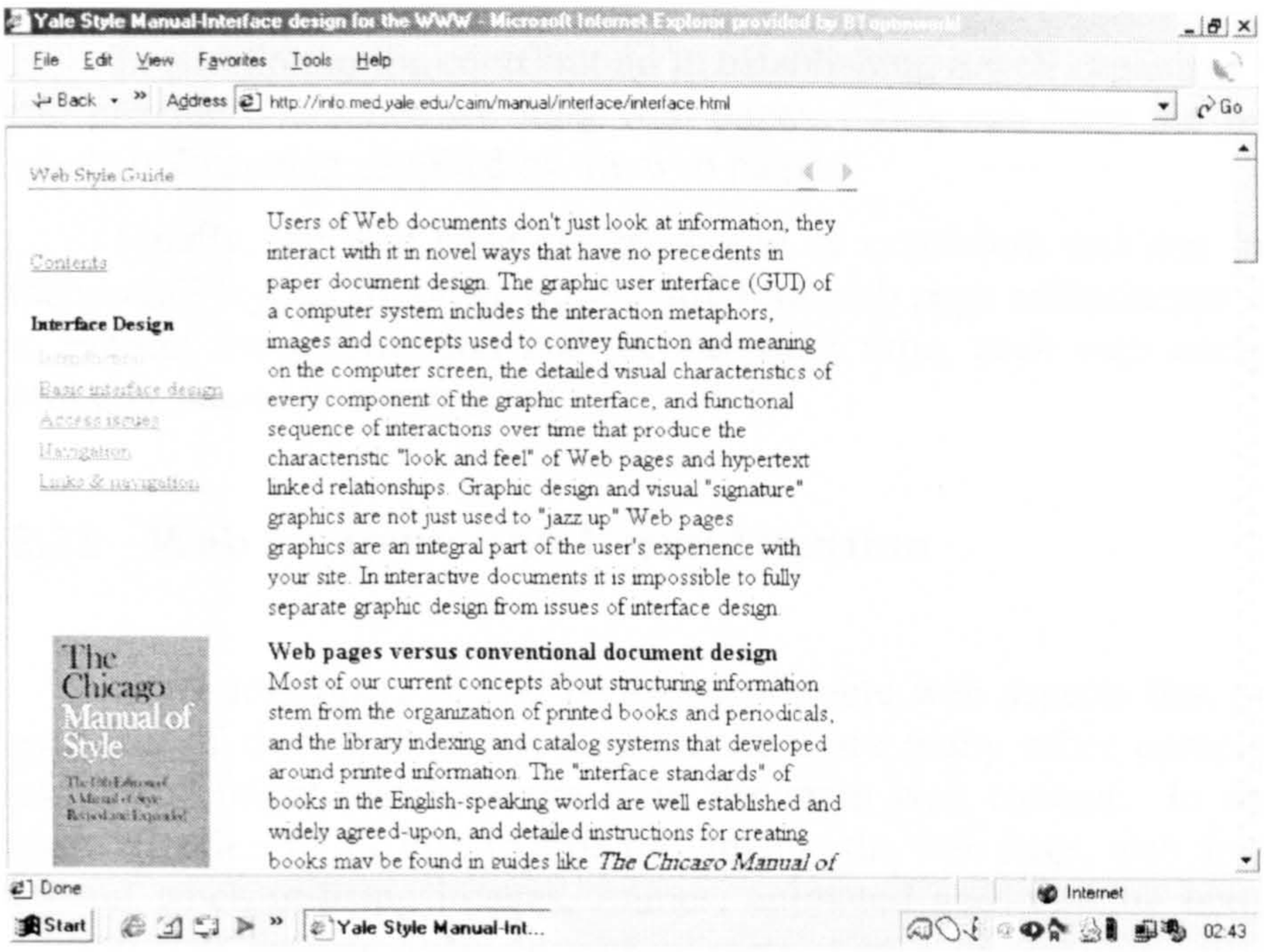


Figure 2.5: Vertical Menu Organisation of a Web Page

There are different ways of implementing menus in the web. The menu presented in Figure 2.5 is a vertical menu presentation, whereas horizontal presentation is illustrated in Figure 2.6. Menus can appear as hyperlinks located at the

beginning of the web document and this type of menu reflects the content of the web page by summarising its headlines.

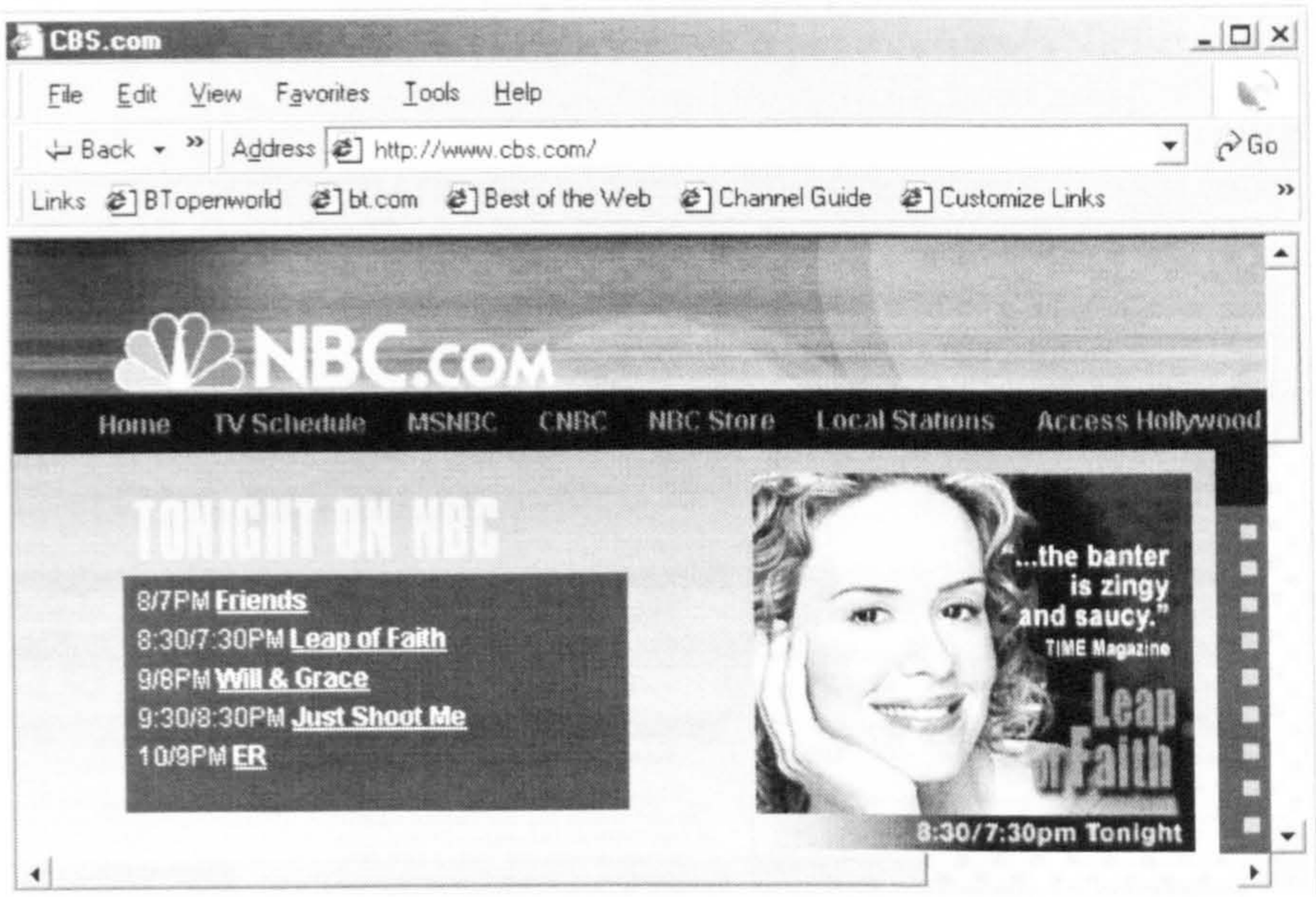


Figure 2.6: Horizontal Menu Organisation

In addition to the contribution in establishing a well organised web page, menus can provide coherence for long web pages. This can keep the users aware of the whole information supplied by the web pages.

Finally, the web page design should be consistent and that involves balancing the overall layout. In other words, different web page components should be selected to present the information and, at the same time, each web component should be properly placed to attract users.

2.11 Web Contents and Users Attention

Web developers should be aware of some web aspects that might decrease the usability of their web pages because there are many other components in the web pages that might distract users from the main web content. In order to attract the users attention to the information provided in the web page, web developers should be careful when utilising images, frames, animated and blinking texts, etc.. The next section will discuss different web experts' views on dealing with web aspects that distract the users' concentration from the web content.

2.11.1 Advertisements

As web pages may be browsed by large number of users many companies want to have a chance to attract advertising revenue. They divide their web pages into

different sections and one of the section is used for advertisements throughout. Users seem to be annoyed by advertisements and have learned to ignore them (Nielsen and Tahir, 2002). But if the developers insist on including the advertisements within their web pages, they should reserve a special places for them, such as a banner area, and never place them next to important web page content (Nielsen and Tahir, 2002).

Figure 2.7: Some Advertisments Areas

Furthermore, the developers should explicitly point to the advertisements if they do not place them in well-known areas often used for advertisements, to let the users easily distinguish them from the main web content (Nielsen and Tahir, 2002). Popup windows advertisements are considered to be one of the worst choices because they keep the users out of the page and distract their attention from the main content (Nielsen and Tahir, 2002).

2.11.2 Images

There is a saying that an image is worth a thousand words, but this might not apply in web pages because the image might take two thousands words of download time (Nielsen, 2000). Different solutions can be applied to solve the image download time problem; some of these solutions are:

- Produce small versions of images (Nielsen, 2000).
- Divide a document containing many images into several documents containing less images (Nielsen, 2000).
- Provide a text alternative for each image.

All the above solutions participate in minimizing the web page's download time, and this has an impact on usability. Users are accustomed to quickly browse web pages and do not like to wait even a few seconds. Developers should remember that web users do not have patience. Animated images and blinking and animated

text, all require more concentration and time from the users in order to recognise the message they intended to convey.

2.11.3 Frames

Frames allow web developers to divide the web page into several sections with each section holding different pieces of information which might be related with one another. Thus, frames can distract the users' attention from the main information. Also, web browsers cannot appropriately print web pages which have more than one frame (Nielsen, 2000); and each frame has its own HTML source code. Search engines have difficulties in indexing the web page containing more than one frame (Nielsen, 2000). To browse each frame users have to scroll each frame separately and if they are not familiar with frame systems that will irritate them.

In order to produce a highly rated web page without annoying the users, developers should exchange the frames for other ways of organising their information. Most of the above difficulties can be overcome by substituting tables for frames or organising the information differently by separating it into several web pages.

Web usability experts' recommendations are aimed at producing web pages which are legible and convenient for the users to navigate. The usability guidelines are based on three aspects: efficiency, effectiveness and user satisfaction. Any component that might break the users' concentration or divert them away from the main information is not recommended by web usability experts.

2.12 Conclusion

This chapter has presented a view of the main features in developing a readable web page. The presentation quantifies the web aspects including the text on the page, colours, fonts, images, page length and width, and the number of links. Quantifying the web aspects measures determines the relationship between different usability guidelines, which give the web developers the ability to select what is best suited for their users. The quantification of web aspects makes it possible to develop an automatic usability measuring approach, described in Chapter 6 Section 6.5.1, that has the ability to evaluate the web pages.

Consequently, later chapters use the web aspects quantification presented above to control the evaluation of web pages, and Chapter 6, Section 6.5.2 describes the utilisation of these measures. Chapter 7 links some of these guidelines and the findings of the usability testing, while the next chapter presents various methods for measuring and evaluating usability.

Chapter Three

Usability Evaluation Methods

3.1 Introduction

An interface is the link between the users and the system or the product. In other words, the interface represents the user interaction with the system in order to achieve some goals. In general the computer interface is what the users see and work with when utilising a specific software product. Computer interfaces are considered to be a vital factor that can be effective or ineffective, efficient or inefficient and that help or hinder the increased utilisation of any products (Hackos and Redish, 1998). Designing an effective interface is a complex task and it does not happen by chance (Sun, 2001); rather it needs a lot of effort to identify a good interface specification. In order to achieve this specification the designers should understand the users of a specific product, the environment the users work in and with, in addition to identifying the targeted tasks for the interface (Hackos and Redish, 1998). Recognising these three interface aspects is considered to be the main factor in accomplishing a good interface. So, a good interface should understand the language the users speak, the users' culture, the users' thoughts, habits and physical capabilities (Hackos and Redish, 1998).

The environment can affect the interface presentation. The interface must be applicable for the screen type and must consider the surrounding area of the users in the design. The tasks embedded in each interface also influence the interface design. These three interface aspects are discussed in detail later in this chapter.

3.2 The Users as an Interface Aspect

The users are considered to be one of the three most important aspects of the interface design. Sklar (2000) and (Lengel, 2002) pointed to the importance of identifying targeted users' needs and profiles in order to increase the efficiency of the interface design. The users' profiles should ideally be a vital description of the users' skills, knowledge, demographic information etc. (Rubin, 1994; Badre, 2002). Generally, as pointed out by Rubin (1994) and Badre (2002), the user profile can be characterised as:

- **Physical history:**

  Users may have some physical challenges which prevent them from easily accomplishing their targeted goals. These challenges can be addressed either individually or by different approaches for different groups.

- **Personal history:**

This contains information about the user, such as the mother language and the other languages he/she might read and understand. The user's personal history might also contain information such as age, gender, left or right handedness and culture. The user's culture has a major effect on the user's tastes, which in turn affect the user's habits in reading, working etc. (Hackos and Redish,1998; Badre, 2002; Lengel, 2002). The user's attitude toward computers is significant, too. Additionally the user's profile might contain the user's learning style, that is the user's habits in learning (Hackos and Redish, 1998; Badre, 2002). Lack of experience with computer systems affects the users' learning to some extent (Nielsen, 1993). For instance, some users prefer to read then do, others might prefer to try and then read, thus learning by doing (Rubin, 1994; Badre, 2002).

- **Education history:**

This refers the user's educational level, such as the highest grade completed, subject studied, primary education etc. The users' education affects their understanding of the interface irrespective of its presentation. The interface appropriate for the primary education users differs from that for users with higher education, as a result of the differences in their knowledge (Hackos and Redish, 1998; Badre, 2002).

- **Computer experience:**

The users' experiences with the computer can be classified into four categories; novice, intermediate, experienced, and expert (Hackos and Redish, 1998; Badre, 2002). These categories are measured on the basis of the total time the users interact with the computer. Habitual use is considered better than infrequent use (Rubin, 1994). Lack of experience when using the computer systems affects the user's learning to some extent (Nielsen, 1993).

- **Occupation history:**

The user's current and past jobs and responsibilities are important since the interface requirements for software will vary to accomplish the goals of dissimilar companies' or departments (Nielsen, 1993; Rubin, 1994; Badre, 2002).

In general, identifying the users' characteristics and needs helps in developing a more effective and efficient design (Hackos and Redish, 1998).

3.3 The Tasks as an Interface Aspect

Each interface design involves several tasks. The main goal, embedded in every interface design, is to reach the highest performance from the users actually using the current interface (Rubin, 1994). However, a web interface has some other goals stated by (Badre, 2002) as:

- Providing consistent information for all users.
- Providing reliable and accurate information.

- Providing a continuous access to the web site or page.
- Providing communication tools for the users.
- Supporting the users with activities that operate within the web environment.
- Providing a searchable web site.

All the above sub-goals work together cooperatively to produce an effective and efficient web page. Their aims are to assist the users.

3.4 The Environment as an Interface Aspect

Users perform their work individually but not in an isolated environment (Hackos and Redish, 1998) since they are influenced by the activities around them, other users working with them and the type of equipment they are using (Hackos and Redish, 1998). Thus, the environment is divided into two basics categories:

- **Physical environment**

The physical environment represents the physical equipment the users work with, screen size and type, mouse, other audio devices, etc. (Hackos and Redish, 1998). Additionally, the space in which the users function, work and interact with the interface is the actual physical environment (Badre, 2002). The users' physical environment affects what the users hear, see and in case of a touch screen. Consequently, they have substantial effects on the success of an interface design (Hackos and Redish, 1998).

- **Cognitive environment**

Despite the possibility that multiple users share the same physical space, each of them has his/her own cognitive environment (Badre, 2002). The cognitive environment is classified into the social and cultural environments. The social environment means that either the user works within a small group or in a larger group; this is relevant as a co-worker or customer might make the use of the interface difficult (Hackos and Redish, 1998). For instance, the user can be frustrated or embarrassed when the interface identifies him/her, to a co-worker, with beep sound as an indication of error.

The cultural environment is not restricted to the ethnic, national or geographic location of the users using the interface design. However, it extends to include three essential elements: religion, language and cultural values (Czinkota and Ronkainen, 1996; Nassif, 2003). The cultural environment, with its three common elements, influences all aspects of users, since it affects verbal and non-verbal communication and the way users behave, think, learn, read, etc. (Nassif, 2003). It is necessary to consider all three cultural elements when creating a multilingual interface. Sun (2001) insisted in localising multilingual web interface to guarantee its cultural elements. Localisation is not limited to translating the content, but includes changing the graphics, colours, symbols, writing direction, time and date format (Sun, 2001).

All the above interface aspects must be considered in order to produce a high standard of interface design. The following section will shed some light on the meaning of usability, usability evaluation and the instruments developed by different researchers to measure usability.

3.5 Usability

The importance of the user interface increased significantly when the personal computer revolution started. Consequently, different slogans have been coined to describe this, but the most important one was the term "user friendly" (Bevan et al. 1991; Nielsen, 1993). All the slogans indicate offering the maximum use of the computer system or product, in other words they aim at making the systems usable. Thus, the usability can be applied to all aspects of a system with which a human can interact starting from the installation to the maintenance (Nielsen, 1993). Several fields have been involved in studying usability such Computer Human Interaction (CHI), Human- Computer Interaction (HCI), User- Centred Design (UCD), Human Machine Interface (HMI), etc. (Nielsen, 1993). These different fields have contributed in making the computer products usable and to the accepted definition of the term usability.

It is important to realise that the usability is not a one-dimensional concept; instead, it contains multiple attributes as described by (Nielsen, 1993; Rubin, 1994; Badre, 2002) learnability, memorability, efficiency, satisfaction and errors. The learnability of the system means it should be easy to adapt to so the users can rapidly establish themselves users of the system. The system should be easy to remember even after a period of time away from it. Efficiency means that regular users should get the maximum productivity of the system after learning it, and they should be satisfied with their results. Also the error rate should be low in order to retain the users' satisfaction.

One of the first usability definitions having been established by (ISO, 1991) for software quality, defines the usability as a set of attributes. However, a more precise usability definition selects certain usability attributes and is stated by (Bevan et al. 1991; ISO 1991; Law and Hvannberg, 2002) as:

"The effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in a particular environment"

This usability definition is considered important since it contains three vital goals every system developed aims to achieve. Another definition of usability is provided by Metrics for Usability Standards in Computing (MUSiC) (Bevan et al. 1991), and is based on the acceptability and learnability. The MUSiC definition is:

> "The ease of use and acceptability of a system or product for a particular class of users carrying out specific tasks in a specific environment; where 'ease of use' affects user performance and satisfaction, and 'acceptability' affects whether or not the product is used"

This definition includes the effectiveness and efficiency, but they are described differently. The effectiveness is represented in the 'ease of use', as described in the MUSiC definition, and the efficiency is the users' performance. The user satisfaction

is shared by both definitions. Therefore, the (ISO, 1991) definition is more descriptive than the other.

3.6 Usability Evaluation

Measuring the effectiveness, efficiency and satisfaction of the interface usability is called the usability evaluation. Usability evaluation is an important interface design process (Ivory, 2001) because of the great benefits of discovering the problems of the design and of allowing a better understanding of the targeted users (Nielsen, 1993).

3.6.1 Usability Evaluation Process

Usability evaluation involved many activities that should be performed as part of the usability evaluation process. This process is made up of several tasks that involve one or more of the interface aspects mentioned in Sections 3.2, 3.3 and 3.4. The usability evaluation process starts with specifying the usability evaluation goals, these being applicable at any stage of the design. The following are typical usability evaluation goals:

- Specifying user interface requirements (Ivory, 2001).
- Confirming or rejecting some designs alternatives (Ivory, 2001; Parush, 2001)
- Identifying specific usability problems (Ivory, 2001).
- Exploring and improving the usability performance (Ivory, 2001; Parush, 2001)

After identifying the usability evaluation and measurement goals. The usability experts should decide whether they want to measure all the usability criteria, which determine the effectiveness, efficiency and users' satisfaction; or they just want to test some of these criteria. For instance, it is reasonable to use usability testing to find out whether the user can successfully use a web site, which includes all the usability criteria measurements. Whereas, the user satisfaction can be used in order to explore and measure satisfaction based on general experience to discover the most convenient web page layout for the users. Thus, the user satisfaction is used in this research to build a general knowledge about the Arabic users' preferences (see Chapter 5). The usability evaluation moves on to perform the following steps presented by (Nielsen, 1993; Ivory, 2001):

- Determine the interface aspects involved in the evaluation. The details of the interface aspects have been discussed earlier in this chapter.
- Select the evaluation method. This is important to effectively achieve the evaluation goals.
- Select the tasks for the evaluation method that are important to determine the targeted users, questions to ask, etc.

is shared by both definitions. Therefore, the (ISO, 1991) definition is more descriptive than the other.

3.6 Usability Evaluation

Measuring the effectiveness, efficiency and satisfaction of the interface usability is called the usability evaluation. Usability evaluation is an important interface design process (Ivory, 2001) because of the great benefits of discovering the problems of the design and of allowing a better understanding of the targeted users (Nielsen, 1993).

3.6.1 Usability Evaluation Process

Usability evaluation involved many activities that should be performed as part of the usability evaluation process. This process is made up of several tasks that involve one or more of the interface aspects mentioned in Sections 3.2, 3.3 and 3.4. The usability evaluation process starts with specifying the usability evaluation goals, these being applicable at any stage of the design. The following are typical usability evaluation goals:

- Specifying user interface requirements (Ivory, 2001).
- Confirming or rejecting some designs alternatives (Ivory, 2001; Parush, 2001)
- Identifying specific usability problems (Ivory, 2001).
- Exploring and improving the usability performance (Ivory, 2001; Parush, 2001)

After identifying the usability evaluation and measurement goals. The usability experts should decide whether they want to measure all the usability criteria, which determine the effectiveness, efficiency and users' satisfaction; or they just want to test some of these criteria. For instance, it is reasonable to use usability testing to find out whether the user can successfully use a web site, which includes all the usability criteria measurements. Whereas, the user satisfaction can be used in order to explore and measure satisfaction based on general experience to discover the most convenient web page layout for the users. Thus, the user satisfaction is used in this research to build a general knowledge about the Arabic users' preferences (see Chapter 5). The usability evaluation moves on to perform the following steps presented by (Nielsen, 1993; Ivory, 2001):

- Determine the interface aspects involved in the evaluation. The details of the interface aspects have been discussed earlier in this chapter.
- Select the evaluation method. This is important to effectively achieve the evaluation goals.
- Select the tasks for the evaluation method that are important to determine the targeted users, questions to ask, etc.

- Design the experiments involved including the determination of the number of the targeted users and the evaluation procedure to be followed.
- Capture usability data from the selected users or other usability metrics by recording them.
- Analyse the data and put it in a form that allows an interpretation of the results.
- Critically evaluate the usability interface in order to recommend improvements
- Iterate. Some data analysis and interpretation needs to be repeated within the usability evaluation process several times to improve the output data and to provide reliable and consistent results.
- Present results. This is the final step in the usability evaluation process. The evaluator presents the results in an understandable way that verifies the usability evaluation goals.

3.6.2 Usability Evaluation Methods

The development of the usability evaluation methods came from the requirement either to recognise the users' preferences, needs and performance or to assess the interface design presentation and interactivity. The usability evaluation methods are classified into two-dimensional approaches, empirical evaluation and analytical evaluation (Parush, 2001; Brinck and Hofer, 2002). In empirical evaluation the users are directly involved in the evaluation process to some extent (Nielsen, 1993; Brinck and Hofer, 2002; Paganelli and Paterno, 2002), whereas in analytical evaluation, various combinations of guidelines, criteria and models are involved in the assessment procedure (Brinck and Hofer, 2002; Paganelli and Paterno, 2002). All the usability evaluation methods are based on testing, inquiry, simulation and inspection methods.

3.6.2.1 Usability Testing Methods

Usability testing is an empirical method, which involve real participants. It is the most fundamental usability evaluation method and the closest to the users. It provides the evaluator and the usability experts with direct and detailed information about how the users can interact with the computer, what problems they face, etc. (Nielsen, 1993; Ivory, 2001). The usability testing procedure involves asking the participants to use the system or perform a set of tasks while a tester or software records the results of their work. The tester then analyses these results and uses them to determine the users' preferences, interface problems and the tasks completion time (Ivory, 2001).

Usability testing is roughly divided into two approaches, the formal test that is conducted as a true experiment to validate some hypotheses, and a less formal test, which is employed to identify deficiencies in the design (Rubin, 1994). The latter test approach is classified into four types according to the targeted problem. The four test

types are exploratory, assessment, validation and comparison (Rubin, 1994). Two important issues should be considered in all the usability testing; reliability and validity (Nielsen, 1993). Reliability is to get the same results as one would if the test were repeated, whereas validity is concerned with verifying that the actual results reflect the usability issues explored in the test (Nielsen, 1993). Seven stages should be conducted to perform the usability testing, and they are:

1. **Test plan** that represents the goals of the test and should be addressed before the usability testing starts (Nielsen, 1993). The plan contains issues like the achievement of the test, the place and the time of the test, the length of the test, the computer support, the software necessary for the test, etc. (Nielsen, 1993; Rubin, 1994). Additionally, it is necessary to set the budget of the test and include it within the test plan. The budget must include the cost of the usability specialist to execute the testing plan, administration, software developers, computers used for the test, usability laboratory and other costs essential to the performance of the test (Nielsen, 1993).

2. **Selecting Participants** or identifying appropriate participants for the user testing. It is one of the most important factors for increasing its reliability (Nielsen, 1993; Rubin, 1994).

3. **Preparation,** which is to develop the test materials that will be used either to communicate with participants, collect data, set the hardware and software requirement for the test (Nielsen, 1993; Rubin, 1994). It is an important activity that should be carried out before the test starts to minimise the number of problems that may influence the test.

4. **Introduction** in order to give the participants a general idea about the usability test goals and to help them focus on their answers (Nielsen, 1993). The introduction should be extended to include some explanation about the system used and if there are any circumstances the participants should take notice of. A reminder of the confidentiality of the data collected and that the target is to improve the interface rather than criticise the participants is useful. Also, it is necessary that the introduction include a notification of the time available for the participants to perform specific task, if the time factor is to be considered in the test.

5. **During the execution of the test** the experimenter should refrain from talking or interacting with the participant unless it is necessary. The interaction could be when the participants are facing a problem that has been observed several times before with other users (Nielsen, 1993; Rubin, 1994). Both positive and negative experimenter reactions affect the participant's attitude and hence, the reliability of the test (Nielsen, 1993; Rubin, 1994). An official experimenter should be appointed if the test has more than one experimenter (Nielsen, 1993).

6. **Debriefing:** it is the reactions and comments from the participants after finishing the test. These comments may be made to justify some answers in the test, assess the questions included in the test and support the experimenters with further comments that might help them identify the difficulties in the test (Nielsen, 1993; Rubin, 1994).

7. **Transforming the data into findings.** This is the final usability testing stage. In this stage the experimenters first analyse the data collected and transform it into recommendations (Rubin, 1994). Two types of analysis can be performed, a preliminary analysis and a comprehensive analysis (Rubin, 1994). The preliminary analysis is a quickly assembled report that is produced immediately and after the test's completion. It highlights the obvious and the common usability testing outlines. But the comprehensive analysis can take two to four weeks after the test performance to produce a detailed and exhaustive report (Rubin, 1994).

These stages can be applied irrespective of the user testing experiment chosen. A user testing experiment has different characteristics according to the targeted goals and participant selection. Many researchers such as (Nielsen, 1993; Nielsen, 1994; Nielsen, 1995; Shneiderman, 1998; Ivory, 2001; Ivory and Hearst, 2001; Law and Hvannberg, 2002; Nielsen et al. 2002; Preece et al. 2002) discussed different user testing experiments:

Thinking aloud protocol: it is to require participants to verbalize their thoughts, feelings, and opinions during the performance of the test. The main advantage of this kind of test is to allow the experimenter to get an idea of the users' mental state during the interaction. **Shadowing** and **Co-discovery** learning are similar to this protocol, however the shadowing method needs an expert user next to the tester to observe the user's behaviour during the test session. The co-discovery learning needs two participants to work cooperatively while the tester observes their interaction. This method is really efficient for discovering the reactions of working participants and is reasonable in resolving working environment interface problems.

Interview method. It is an extension to the thinking aloud protocol; however, in this method the tester asks the participants specific questions about the design. It is obvious that the interview method is more focused and directs the users to certain design issues. This method also enables the experimenters to better understand the participants' mental model.

The Coaching method. It allows the users to interact with the testers and ask them question regarding the interface tested. It is considered to be the reverse of the interview method. The testers can learn how the users interact with the system and they can directly identify the problems that face the users.

The Teaching method which starts by allowing the participants to interact with the system first to develop experience of the system. Then these participants teach the novice users how to use the system in addition to solving any problems they might face. By using this method the experimenters can assess the learnability of the system.

Remote testing. It is to give the participants questionnaire to answer, one which involves questions to achieve specific goals. The main difference that distinguishes this method from the previous ones is that the participants and the experimenters are not co-located during the test.

Observing method, in this method the tester visits the users and observes them working without interfering with them; the tester can also videotape them for further analysis. This is the simplest method of testing since there is no need for any test preparations in advance. However, if the users are videotaped during the test and the tester and the participants review the videotape together, the method will also be

Retrospective testing. This kind of testing can be used with any other method to support the experimenter and the test goals.

Performance measurement. It has the goal of capturing quantitative data about participants' performance while they accomplish specified tasks. This type of usability testing is run in a usability lab, and the experimenter can interact with the participants during the test. Sometimes this method can be supported with other methods, such as retrospective testing mentioned earlier.

In all of the user testing methods presented above, it is obvious that they are really close to the users. In other words, the experimenters learn about the users' feelings, attitudes, preferences and, most important of all, their mental model (Nielsen, 1993; Rubin, 1994). This research relies on a user testing method to identify Arabic and English users' preferences.

3.6.2.2 Inquiry Method

This method is close to the usability testing method because it requires feedback from the users (Ivory, 2001; Ivory and Hearst, 2001). However, it does not focus on specific tasks or measuring performance (Ivory, 2001; Ivory and Hearst, 2001). Different types of inquiry methods exist and are based on the evaluator's interaction with the users. In some types the evaluator accesses users' reported information rather than having a direct interaction; for instance, users feedback, screen snapshots, log-files and self-reporting logs. In the last two types the users write their actions, observations and comments using either paper and pencil or computer forms (Ivory, 2001; Ivory and Hearst, 2001). The evaluators interact directly with the users in contextual enquiries and focus groups. The contextual enquiry is usually a long-term method, which may lasts for one year (Ivory, 2001; Ivory and Hearst, 2001). The focus group is an efficient method to use in order to improve the usability of a particular product for future release (Ivory, 2001; Ivory and Hearst, 2001).

3.6.2.3 Inspection Method

The usability inspection method is an evaluation method that relies on an evaluator or a team of evaluators who examine the interface design. All forms of this method are based on the adoption of a group of guidelines and design principles, but they differ in how they are processed and how to apply them. Also, the inspection method types are varied according to the goals of the inspection as expressed by different researchers (Nielsen, 1993; Nielsen, 1994; Nielsen, 1995; Shneiderman, 1998; Ivory, 2001; Ivory and Hearst, 2001; Parush, 2001; Law and Hvannberg, 2002; Preece et al. 2002; Rohn et al. 2002):

Guideline review; this is where the evaluators verify in advance a comprehensive and sometimes large number of established usability guidelines. Interface design experts and researchers will have established these guidelines and standards. The guideline review method has a limitation since the adoption of a wide range of usability guidelines may lead to conflicts between them, a lack of empirical evidence and they may be too general to apply to specific interface cases.

Cognitive walkthrough: it involves one or more interface design evaluators, who explore interface design or prototype by focusing on a specific set of tasks to assess their understandability and learnability. During the assessment the evaluators simulate the users' behaviours and by the end they construct a detailed success or failure story about each task attempted. This method requires intensive documentation to present the acceptable and unacceptable walkthrough task sessions. If the test involves the user, developers and evaluators in inspecting the interface, then it is called the pluralistic walkthrough.

Heuristic evaluation. This requires a small group of evaluators. Each member in the group independently evaluates an interface using a list of heuristics. The usability rating outcomes from all the evaluators are aggregated and merged to determine reasonable usability rate for each aspects tested and to resolve potential usability problems. It is an effective method to improve the interface or product usability; also it is an effective way to educate in colleges and to develop processes.

Consistency inspection. It is to investigate the differences and the consistency between different interface designs. The investigation is performed by a team of interface design evaluators who inspect one interface at a time.

Other inspections methods exist, such as card sorting, storyboarding, perspective-based inspection, feature inspection, formal usability inspection and standards inspection mentioned in different studies (Nielsen, 1993; Nielsen and Landauer, 1993; Nielsen, 1994; Nielsen, 1995; Sawyer et al. 1996; Ivory, 2001; Ivory and Hearst, 2001; Rohn et al. 2002).

3.6.2.4 Simulation Method

This method is performed using a model of a user and/or an interface design. A computer program mimics the user's interaction with the interface and reports the results (Hilbert and Redmiles, 2000; Chi et al. 2001; Ivory, 2001; Ivory and Hearst, 2001; Brinck and Hofer, 2002). The report produced includes the performance and the interface operation such as mouse clicks, different functions selected from the interface, etc.. The evaluator can run the simulation with different parameters and learn lessons about the user's reaction to a specific user interface. This method is widely used in automatic usability evaluation and has been used in this research. The discussion of the webmaster's feedback simulation is presented in Chapter 6, Section 6.5.5.

3.7 Usability Testing for Cultural differences

As mentioned earlier in the usability testing method Section 3.6.2.1, usability testing is considered to be the closest method to the users, in the sense that it reflects the users' thoughts, feelings, working habits and preferences (Nielsen, 1993; Ivory, 2001; Ivory and Hearst, 2001). Therefore, usability testing techniques have been widely used to investigate new guidelines, to confirm and improve existing guidelines (Nielsen, 2000), and to study cultural differences.

The users and their cultures affect user understanding and performance; therefore these aspects have been considered by many researchers (Russo and Boor, 1993; Marcus, 2001; Sun, 2001; Walton et al. 2002; Laarni, 2003; Walton and Vukovic, 2003). Two types of user testing have been used in Walton and Vukovic's (2003) study to investigate how South African novice and experienced students understand some common interface and navigational conventions on the web. Walton and Vukovic combined the think-aloud protocol and the teaching method and ran them over 20 students. Additionally, they considered cultural differences between South Africa and Western countries in terms of the specified set of interface and navigational conventions. The findings showed significant differences between the two cultures in visualising the information hierarchies (Walton et al. 2002; Walton and Vukovic, 2003). In other words, the differences are represented in the organisation and visual representation of information structures (Walton et al. 2002; Walton and Vukovic, 2003).

The internationalisation of web pages refers to producing them in many languages, in other words, developing multilingual web page versions without considering the interface of cultural aspects (Russo and Boor, 1993). But it is not enough to localise a web page by simply translating the content with some modifications to the data or number formats, without localising the interface design layout to the users' preferences and needs as well. Two localisation aspects must be addressed when developing multilingual web pages. The first is the language translation, measurements, currency, addresses and so on, while the second is the aesthetic layout of the interface design which reflects cultural criteria (Sun, 2001).

In order to create usable multilingual web pages that are effective, efficient and satisfy the users' needs, (Sun, 2001) conducted user testing on three international communities from China, Germany and Brazil. Sun (2001) focused his study on examining the four major cultural categories of language, visuals, colours and page layout. The first category addresses the surface level of the localisation, while the other three categories are related to the secondary localisation aspects. The main aim of his study was to learn how the four cultural categories might affect web usability. He interviewed the target users about their experiences using the localised version of both the Lotus and Adobe web pages. It appeared that different users tried to apply their own cultural preferences to evaluate the design of the web pages. For instance, the Brazilians like to see vibrant colours and lively pictures on their local web pages, while the Chinese prefer to see one of the common Chinese flowers on their web pages. The Germans like the links in the navigation bar to be alphabetically organised. The main implication from Sun's study is that consideration of the cultural categories can increase the web page's usability.

The influence of culture can reasonably be expected to extend from usability to usability testing results, especially if the usability test is performed in languages different from the participants' mother tongues. Understanding the participants' culture, which is part of the users' profile, increases the efficiency of user testing (Yeo, 1998). Other user testing conducted by (Goonetilleke et al. 2001) focused on four factors: interaction efficiency, trust and safety, information content and access, and input-output; in order to simplify usability questionnaires, enhance the interface and reduce the user testing time. The interaction experience of sixty four participants were involved in (Goonetilleke et al. 2001) study. They divided the participants into two groups of thirty two. Each group had to interact with either specified airline or retail sites and then fill in a questionnaire expressing their opinions. The overall

results showed that a significant difference exists between the retail and airline sites with reference the four factors tested. Goonetilleke, Duffy and Jacques concluded "this finding has important implications in terms of website design as well as usability testing. The results may be used to simplify usability questionnaires".

Remote usability testing has been used for collecting interaction data and studying the cultural effects on user interface design. Lee (1999) conducted cross-cultural remote usability testing of a total of 172 respondents from 15 countries, although he only compared Korea, Japan and United States of America. The questionnaire answers and the interaction were saved and analysed to identify relationships between the participants' interaction and their cultural characteristics. The evidence of the (Lee, 1999) study showed some differences among the three cultures selected. For instance, one of the questions he asked was whether the colour of the power button in the remote control was 'red' or 'green'. A high percentage, (91%) and (87.5%) of the Koreans and Japanese respectively chose 'red' while (89.4%) of the Americans chose green.

Culture is an important ingredient in identifying information technology specifications and it has effects on both the use and the interface (Little et al. 2000; Marcus and Gould, 2000; Marcus, 2001). It is clear from the evidence of previous studies that culture plays an important role in determining reasonable usability guidelines for the targeted culture in addition to its influences on the users' preferences, thoughts, beliefs etc.. Different types of user testing have been used effectively to identify cultural effects on user interface design. Therefore, this research is included utilising usability testing to identify Arabic users' preferences and then establish usability guidelines appropriate for assessing Arabic web pages. Additionally, usability testing is used to validate some of the current usability guidelines for English web pages.

3.8 Automated Web Usability Evaluation

The tremendous number of existing web pages, developed both by professionals and by people who have little design' experience (Ivory, 2003), have created a variety of web design styles. These styles differ according to the languages, ages, cultures etc. represented. In order to improve web page design existing pages need to be evaluated and, because there are so many, they need to be assessed automatically and with reference to the users' preferences and beliefs. It is increasingly necessary for web designers and usability specialists to take advantage of automated usability tools and approaches as part of the overall usability process (Brinck and Hofer, 2002). Automated web usability approaches can save the usability specialists and designers money and time (Ivory, 2001; Ivory and Hearst, 2001; Brinck and Hofer, 2002; Ivory and Hearst, 2002) and improve site design systematically. In other words, automated web usability assessment tools can discover various error types consistently, increase the quality of design by enabling comparisons between various designs and produce reliable usability standards (Ivory, 2001; Ivory and Hearst, 2001; Brinck and Hofer, 2002; Ivory and Hearst, 2002).

Currently there is a variety of methods available to automatically assess web pages. As mentioned earlier some of them are concerned with identifying problems in web design that might discourage users from performing certain tasks, whereas others

compare the web design with existing highly rated designs, and use them as a basis for evaluating other web pages. Quantifying different web aspects is also an evaluation technique, but it is unreasonable to rely on raw numbers or quantities as a measure of quality, particularly when evaluating a dynamic environment like the web.

Web Event-logging Tool (WET) (Scholtz et al. 1998; Etgen and Cantor, 1999) automatically collects usability data about users by remotely tracking their interactions with the page, and recording events. WET records sophisticated events, logging things such as mouse clicks, changes and motion as well as page load (Hong et al. 2001; Paganelli and Paterno, 2002). It utilises the existence of the event handling capabilities built into Netscape Communicator and Microsoft IE (Etgen and Cantor, 1999; Hong et al. 2001). WET has three major disadvantages; firstly it is limited to the two web browsers mentioned. Secondly, the tester must restrict WET's information gathering to particular events or objects on a web page to avoid the data becoming overwhelming, and lastly the tester has to analyse the collected data manually (Etgen and Cantor, 1999; Hong et al. 2001; Paganelli and Paterno, 2002).

However, other authors take a different approach which avoids these problems. WebQuilt (Hong et al. 2001) works independently as a proxy server without relying on any web browser. It monitors the user's movements until he/she locates a specific piece of information (Hong et al. 2001). It is a visualized system that can be run locally and remotely. The visualisation indicates the most commonly traversed path of the users through a web site for a given task combined with the time spent on each path. This visualisation also differentiates the longer, in terms of time, from the shorter, associating this with the optimal path envisaged by the web designer. In order to collect the users' data WebQuilt utilises the proxy approach which overcomes many of the server-side or client-side problems (Hong et al. 2001).

Web Remote USer INterface Evaluator (WebRemUSINE) (Paganelli and Paterno, 2002) supports the tasks of retrieving, navigating and accessing information, based on links to remote pages. The users are asked explicitly about the task they want to perform. Once this information has been collected the system implicitly observes the user's interactions (Paganelli and Paterno, 2002). WebRemUSINE replaces the proxy server logging method with java applets that activate at the beginning of the test to log user interactions, and then provides all the data to a servlet on the server-side. All the data collected is saved into file to be evaluated (Paganelli and Paterno, 2002). WebRemUSINE (Paganelli and Paterno, 2002) presents an intelligent way to determine failures of the design being tested in that it deters or prevents the users from getting the necessary information. Paganelli and Paterno (2002) associate the users' actions with the targeted task to test the efficiency of the web application design.

Other approaches assess the web design based on predetermined guidelines or criteria. These guidelines might be empirically revised before their adoptation. WebTango, (Ivory, 2001; Ivory and Hearst, 2002) one of the current approaches, adopts some guidelines collected from and driven by previously highly rated web pages and then statistically processes them to produce an impartial evaluation tool. WebTango's procedure is to calculate web metrics collected from the HTML source code and then to generate an appropriate usability assessment. Ivory introduced WebTango (Ivory, 2001) that can evaluate the web page based on a statistical model. This model is constructed from the highly rated web pages, seeks some guidelines from them, and utilises them in the web page evaluation. Each web page is analysed to compute twelve quantitative measures (Ivory, 2001; Ivory and Hearst, 2002).

These measures have to do with page composition, layout, amount of information, and size such as the number of words, links and colours (Ivory, 2001; Ivory and Hearst, 2002). The limitation of this approach is that it does not provide a customised evaluation because the presentation of information might differ as a result of changes of the designs' objective circumstances. For instance, using different frames is not recommended (Lynch and Horton, 2001), but certain designs need frames in order to organise the data.

Most existing automatic usability evaluation tools such as Web Static Analyzer Tool (WebSat) (Scholtz et al. 1998) rate the web design through checking the validity of some web usability principles in the web applications and these tools are initiated by the user; whereas, WebTango (Ivory, 2001) can be considered to be a benchmarking evaluation tool since it evaluates the web application based on previously high rated web pages. Neither of them is designed to modify the usability principles or customise them to the users' preferences, but they can compute a set of quantitative web metrics.

LIFT is software developed by UsableNet Inc. (LIFT, 2000). This software allows web designers and owners to test the accessibility and usability of their web sites (W3Consortium, 2003). It is implemented to work with web editors such as Dreamweaver, GoLive, FrontPage, BB Edit. It runs locally on MacOS (W3Consortium, 2003). This software is integrated to detect the web developers' mistakes from the initial development stage. However, it is better to provide automatic web usability evaluation that is not restricted to certain software or specific platforms.

Therefore, the need for an alternative software model that automatically utilises the data collected from web pages and associates it with the users' taste becomes a focal point. Furthermore, constructing a tool based on specific usability principles, which are well established by many research studies will contribute to identifying outdated principles and the need to update others.

By utilizing different software models, which support usability evaluation, this research tries to integrate them to provide a reasonable solution for automatically updating the current web usability principles based on experts' and researchers' recommendations and to customise them to the users' preferences. Also, by considering the multilingual aspect in web design, this research will extend web usability guidelines, by user satisfaction, to Arabic web pages as well, since there are no specific Arabic web usability guidelines at present. These guidelines will be used as the basis for Arabic web page evaluation, and when combined with English web pages guidelines, can be generalized for languages which follow the same writing direction with similar scripts.

3.9 Conclusion

It is impossible to create a universal interface design that is applicable to all cultures and societies. The development of a highly usable interface design requires strong background in identifying the three interface aspects, targeted users, interface design goals and the interface environment. Several usability evaluation methods have been established to minimise the problems that can occur after developing an

interface. The most common usability evaluation methods used are the heuristic evaluation and the usability testing because of their effectiveness and reliability.

Several of these usability evaluation methods have been automated to increase productivity. Different automatic web usability evaluation tools have been established to solve the diversity web interface problems. However, few of them have been characterised by the intelligence necessary for the next generation of automated web usability. To customise the evaluation according to the webmaster's needs, which will reflect the targeted users' desires, intelligence must be added to the architecture of the automated web usability tool. The following chapter discusses several intelligent approaches appropriate for introducing intelligence into automated web usability evaluation.

Chapter Four

The Agent Technology

4.1 Introduction

The previous chapter introduced different methods for usability evaluation. Lack of intelligence was identified as one of the weaknesses of the majority of existing tools. In this chapter different intelligent and learning algorithms such as neural network, fuzzy logic, genetic algorithms and reinforcement learning will be discussed. In order to investigate such techniques it is important to develop an impartial system in which they can be applied and for this purpose an agent technique is used. Creating impartial agent systems requires sensing the input, processing the problem, producing appropriate actions, and adjusting future actions. The agent can contain one or more intelligent or learning algorithms depending on its goals.

The dictionary definition of an agent is: "agent n. 1- One that acts or has the power to act. 2- One that acts for or represent another. 3- A means of doing something. 4- Something that causes a change." (Dictionary 1994). While the agent's definition in the (Allen 1984) is "1- One who acts for business, etc. 2- One who or that which exerts power or produces effects." The combination of these definitions can be applied to software agents, hence "software agents are active, independent components" (Millier 1997). Most software agents are implemented to act on behalf of a user or other programmes to perform tasks (Maes 1994; Terveen and Murray 1996). An example of this kind of agent is the finance agent that buys a specific number of shares when the stock price falls below a specific level (Terveen and Murray 1996).

In other contexts, software agents are designed to act as a user in order to help execute some task or operation (Franklin and Graesser 1996) that would be hard to achieve for a human user because it would be too repetitive, tedious, or time consuming (Maes and Wexelblat 1996; Terveen and Murray 1996). In doing this, the agent must have some knowledge of the user's preferences, goals, habits and desires (Terveen and Murray 1996). A good example of an agent that substitutes for the user doing a routine hard to organize and dull to process is the meeting agent *visitorbot* (Kautz, Selman et al. 1994). It uses information about a group of people that want to arrange the meeting and suggests a schedule that satisfies as many constraints as possible. The collection of knowledge accessible by the agent should be based on user's preferences, interests, availability and agenda (Kautz, Selman et al. 1994).

At present there is a variety of agents that assist users in different ways. Each one of them tries to empower users by providing valuable support to enable them to work more effectively in the rich and ever-changing world of electronic communication and information (Terveen and Murray 1996).

4.2 Agent Definition

The vast growth and improvement in web agents has resulted in their definition varying from one researcher or developer to another with each one trying to express his or her own view of an agent; furthermore the definitions differ according to what the agent is doing (Nwana 1996). Therefore, it is hard to identify one definition for all kinds of agent.

In general, Nwana (1996) defined an agent as either software or/and hardware component that acts on behalf of its user. Therefore, a software agent is a particular type of agent (Maes and Wexelblat 1996). Many other definitions of software agent exist; (Yap and Keong 1999) present the notion of it as a programme that performs a set of tasks, is designed for the benefit of one or more users and lives in general purpose computer networks, computer operating systems, databases, and other similar environments (Franklin and Graesser 1996). This definition is the basic foundation of software agents and is discussed by many researchers and developers such as (Franklin and Graesser 1996; Nwana 1996; Terveen and Murray 1996; Maes 1997).

However, a more advanced agent is a self-determining, computational based programme or a group of programmes that act in parallel with the user but without his/her intervention (Lieberman 1997). Lieberman's definition points to one of the most important agent characteristics; one which distinguishes it from conventional software programmes. Self-determination includes both self-reaction and autonomy. These two characteristics help the agent to decide its goals independently and to plan toward the functions that it was designed for.

The agent has sensors to learn about the environment and the user's needs; in addition it has effectors to perform actions in order to fulfil its goals (Maes and Wexelblat 1996; Maes 1997). Any changes that occur in the environment can affect its behaviour, cause the rescheduling of its agenda or the use of its senses in the future (Franklin and Graesser 1996). Lieberman (1997) adds that a software agent remains active even when the computer is turned off. For example, a web agent must be able to operate even when the user is not connected to the Internet or has no process on the WWW. So an information filtering agent continues to look for any new article that mentions a specific word or phrase specified by its user (Maes 1994).

Some definitions have been proposed to cover the capabilities of software agents. Coen (1994) represents this idea when he defins software agents as autonomous programmes that engage in dialogue, negotiation and the coordinated transfer of information. Additionally, autonomous agents are programmes that initiate communication, set up meetings, monitor events, perform tasks and locate information from different databases (Kautz, Selman et al. 1994; Maes, Shneiderman et al. 1997).

Software agents range from simple macros or subroutines where the user enters a few parameters to more complex and intelligent programmes that make decisions and perform complicated tasks (Genesereth 1994; Kautz, Selman et al. 1994; Do, March et al. 1996; Millier 1997; Das, Caglayan et al. 1998). Usually, software agents are constructed based upon a set of rules they follow when first initiated (Do, March et al. 1996). However, this set can be expanded as the software agents learn more about their users' working habits, and observe their environments. Software agents are capable of adapting their behaviour and updating their performance to meet these

changes. That means software agents understand users' needs and the requirements of their environment. This allows the software agent's reasoning capabilities to improve (Do, March et al. 1996). The level of software agents' intelligence varies from one agent to another according to their ability to adapt to different circumstances, and to a number of other characteristics they possess (Millier 1997).

It is obvious from the definitions of software agent that some are likely to exhibit different characteristics of an agent in different degrees (Millier 1997). The major characteristics of agents have been discussed above. The next section focuses on some of important characteristics in more detail.

4.3 Agent Characteristics

There are sets of criteria which can measure the quality of software programmes in general; criteria such as correctness, completeness, efficiency, and safety (Friedman and Nissenbaum 1996). However, agency, intelligence, competence and trust are further specific criteria that relate to software agents (Do, March et al. 1996; Friedman and Nissenbaum 1996).

Software agents have some characteristics which make them distinguishable from the ordinary programmes. They have the additional dimensions of autonomy, mobility, agency, independent of their owners' or users' preferences, and are faster, stronger and more intelligent than conventional software programmes (Do, March et al. 1996). It is important to introduce the characteristics of agents that differentiate them from all types of software and applicational programmes (Murch and Johnson 1999). Also, it is necessary to consider characteristics of agents relevant to the multilingual automated web usability evaluation agent, developed by this research, because they are valuable when performing a customized and customisable evaluation of web pages. The agent is characterised by one or more of the following:

- **Autonomy:** The ability of the agent to act autonomously without the user, or any other programme or software intervention controlling its internal state and actions (Knapik and Johnson 1998; Murch and Johnson 1999). It is considered the key characteristics of the software agent's behaviour (Millier 1997). This means that the software agent considers its goals, decides how to achieve them and acts upon these decisions without oversight or management by any other entities (Das, Caglayan et al. 1998; Barber and Martin 2001). The software agent is responsible for determining how to pursue its goals (Wooldridge and Jennings 1995; Das, Caglayan et al. 1998; Barber and Martin 2001). This feature has been considered in the multilingual automated web usability evaluation agent. Thus, it is independently willing to perform the analysis and evaluation without any intervention in its internal actions (see Figure 6.1).
- **Adaptability:** The agent should be able to work on multiple platforms and it should learn from its user and adjust itself to the user's working habits, previous experience, and preferences. Therefore, a characteristic behaviour of an agent is to adapt to the modification of its environment and changes, based on user's methods without any direct input from him/her or the owner

(Franklin and Graesser 1996; Murch and Johnson 1999). This feature can be described as personaliseability and has been considered by many researchers (Maes and Wexelblat 1996; Seo and Zhang 2000). This is also considered in the multilingual automated web usability evaluation agent since the agent is required to customise its evaluation to the webmaster's needs (see Chapter 6 Section 6.5.3).

- **Collaboration:** This refers to the agent's ability to share, which allows it to cooperate with other agents or their users in the performance of some computer-based tasks (Rich 1996; Moran, Cheyer et al. 1997).

- **Communicativeness:** In order for an agent to learn from both the user and the environment, it should have the ability to communicate with repositories to get required information (Cesta and D'Aloisi 1996; Franklin and Graesser 1996; Knapik and Johnson 1998; Murch and Johnson 1999). The agent may carry out a discourse with user or share other agents' agendas and carry out its task in the way they prefer.

- **Cooperativeness:** This aspect is exhibited when the user specifies actions that need to be performed by the agent. The agent responds in being willing to provide its owner or the user with the results. The agent-oriented system has to know the needs of the agent's owner or user, and then plans for and decides its next action (Cesta and D'Aloisi 1996). This feature is exhibited in the multilingual automated web usability evaluation agent when it shows its ability to evaluate the web pages assigned by its owner (the webmaster) and considers the webmaster's feedback in customising the next evaluation.

- **Reactiveness:** The agent needs the ability to decide when to start or stop in a timely fashion responding to changes in the environment (Franklin and Graesser 1996; Knapik and Johnson 1998; Murch and Johnson 1999).

- **Mobility and flexibility:** An agent should be flexible enough to transport itself from one machine to another and have the ability to migrate through a network to find information in order to achieve its goals based on internal decisions (Do, March et al. 1996; Franklin and Graesser 1996; Moran, Cheyer et al. 1997; Murch and Johnson 1999). Also, it may have the ability to execute in more than one processor, in order to support the execution of other programmes or software agents that need help in performing their delegated tasks (Knapik and Johnson 1998). It must act as a mediator in performing tasks and dealing with different aspects in the Internet, such as nodes and packets transformation and their security measures in the network (Do, March et al. 1996; Knapik and Johnson 1998). The requested data, which is retrieved by mobile agent, either returns to the origin of the mobile agent or to some other place that requested the information (Do, March et al. 1996).

- **Distributedness and transparency:** In order to accomplish complex tasks, an agent might interact with other agents in the network to share the completion of the task in transparent manner (Cesta and D'Aloisi 1996; Moran, Cheyer et al. 1997; Murch and Johnson 1999).
- **Rationality:** The agent must act in a manner that fulfils its goals instead of preventing it from achieving them. In other words, the agent must solve as many problems as it can and be able to deal with errors and incomplete data (Murch and Johnson 1999). This characteristic has been used in the multilingual automated web usability evaluation agent when analysing the web pages and computing the web metrics necessary for their evaluation (see Chapter 6, Section 6.5.2).

One of the most important features that differentiate software agents from conventional software programmes is temporal continuity, which means that they are continuously running programmes and do not terminate after getting the result (Do, March et al. 1996; Franklin and Graesser 1996; Lieberman 1997; Murch and Johnson 1999), unless they are terminated by their owner. The above are some of the important characteristics of agents, but their presence varies from one agent to another. For instance, *Apt Decision* agent acts as an estate agent that helps users in selecting their favourite apartment (Shearin and Lieberman 2001). This agent learns the users' preferences as the users criticise the options offered by the agent. This agent is immobile. However, the buyer/seller agent presented by (Lynden and Rana 2002) is used to buy items. First the agent determines which item to buy and then it roams the network to find an appropriate purchase. This kind of agent needs to have a mobile feature in order to achieve its goal.

4.4 The Importance of Agents

Software agents are programmes that assist people and act on their behalf (Lange and Oshima 1998; Yap and Keong 1999). Agents perform tasks delegated by the users who are unable or unwilling to do them, because these tasks are repetitive, tedious, dull or too overwhelming to perform (Maes 1997). Agents have been successful in many areas, stated by (Knapik and Johnson 1998):

- Searching a large problem space and problem reduction. This is used to solve games problems such as chess.
- Logical reasoning, deduction, proven theorem, and the analysis and maintenance of databases.
- Language understanding such as reading text, translating and word understanding.
- Learning using intelligent techniques and algorithms that emulate normal feedback. According to feedback the algorithms will tune the agent behaviour to produce the expected output.
- Knowledge-based and expert systems which convert human expertise within a domain of knowledge to forms usable by others.

Robotic and vision, operating systems, tools, security, personal shoppers, librarians and many other applications for agents have been discussed by different authors (Wayner 1995; Knapik and Johnson 1998; Lange and Oshima 1998; Murch and Johnson 1999).

4.5 Interface Agent

The user interface can be considered as a medium to transfer messages between the user and the computer system or programme that the user operates on. This interface plays an important role in the relationship between the users and the computer system (Millier 1997). In practice, the computer interface could represent a major obstacle that could prevent the user from utilising the system; alternatively, a user-friendly interface attracts people and makes the computer system or application programme popular and likely to be used. The inventors and developers of the computer systems focus on promoting utilization for many existing computer schemes. They implement a personal assistant which interacts and collaborates with users to support them in performing tasks in a variety of ways (Koda and Maes 1996). One of the most important services that supports personal assistants is the interface agent, which completely hides sophisticated tasks from both the user and the application developer (Maes 1995; Cesta and D'Aloisi 1996; Nwana 1996) and suggests alternative solutions as well.

A typical job for the interface agent as (Nwana 1996) stated is to assist a user in learning a particular application (Lashkari, Metral et al. 1998). Nwana (1996) summarize the working methodology for an interface agent as: firstly, the agent observes and monitors the user's actions; secondly, it learns a new shortcut or faster technique to do the task; finally, it recommends an improved way to perform the user's task. The interface agent does not prevent the user from choosing their own way to accomplish a task (Maes 1997).

However, (Bayardo, Bohrer et al. 1998) declare the interface agent to be one that replies to the user's queries, typed and displayed in a natural language. *InfoSleuth* is project for filtering requested information composed of several agents, of which one is a user agent responsible for the interface with the user (Bayardo, Bohrer et al. 1998), whereas, (Cohen, Cheyer et al. 1998)'s interface agent is integrated into the open agent architecture. It is obvious that this is a specific class of interface agent and the most simple. Etizion and Weld (1998)'s *Softbot* is an interface agent that has further expressiveness in its interface. It accurately identifies the information requested from a hint of what the human wants, then takes responsibility for how and when to accomplish the request in a single, expressive and consistent interface (Etizion and Weld 1998). In other words, *Softbot* can convert the user's high-level goals into more detailed goals (Knapik and Johnson 1998).

Other interface agents save a user profile used during the execution of the application to broaden the agent's knowledge in recognising the user's specific goals. The *DIA* interface agent assists a programme debugger used in reproducing the state of the programme to be debugged as well as having experience in handling similar tasks in an effective manner (Lieberman 1995; Goecks and Shavlik 2000; Seo and Zhang 2000). In other words, *DIA* aims to help the user to automate routine tasks with a certain degree of success; furthermore it attempts to process unusual tasks

when presented with unpredictable pieces of user input that are rarely repeated (Fujishima 1997).

The interaction between the user and the agent in the system can be via conventional input entry such as keyboard and mouse. Additionally, speech, handwriting, pen-based gestures and many other input tools are possible. However, high quality interface agents, which possess special skills and have rich experience in navigating the WWW, work without any direct input from the user (Beale and Wood 1994). In other words, they act autonomously to reach their goals without interaction with the users (Beale and Wood 1994).

Finally the interface agent can be defined as an invisible programme that continuously executes as a background process, and can observe the user input and output from the system. Therefore, the interface agent can learn from its user and reflect on what the user sees and does; it is not necessarily a one to one exchange with the user actions. The user does not send orders to interface agents; rather if the interface agent meets its requirements, it determines how and when to achieve the user goals (Fujishima 1997; Lieberman 1997; Brown, Jr. et al. 1998; Etizion and Weld 1998). Different ways for explaining how the agent might learn from the users will be discussed later in this chapter.

4.5.1 Classification of the Interface Agents

There are different ways of classifying interface agents, depending on their functions or their properties (Beale and Wood 1994):

- **User agents:** are self-customising pieces of software. They can adapt and learn their users' working habits in order to personalise their behaviour to their users' preferences. The performance of the user agent will be unique for each individual user, such as the e-mail classifying agent.
- **Guide Agents:** act as tutors supplying the users with knowledge they need to improve their skills. These agents will also enable the users to improve their expertise as is the case, for instance, with Microsoft agents. The agent in this case assumes that the user has no background knowledge about the topic, and it takes an active role in leading the user through the desired task.
- **Cooperative Agents:** they are user assistants. They adopt certain tasks to work with and act as an expert. They provide the users with alternative views or additional supporting and relevant information. For example, the *CAD* helper supports the construction of engineering drawings in providing the start, midpoint and tangents to lines.

It is obvious that with the cooperative agent, the users have some idea about the topic but need either more information or another way to handle their tasks. The cooperative agents criticise the users' work and suggest better alternatives (Beale and Wood 1994).

All of these interface agents are involved in personal user assistance, but each of them has different degrees of interaction with the user and the interaction involves some degree of intelligence in the interface.

4.5.2 Aspects of Intelligent Interface Agents

The intelligence of interface agents can be measured by the degree of understanding of the users' needs they have and by their ability to sense changes in the environment. This intelligence can affect the interface agent's behaviour and interactions with the users. The behaviour that must be exhibited in an interface agent for it to be said to possess a specific degree of intelligence has been discussed by different authors (Beale and Wood 1994; Brown, Jr. et al. 1998; Amant 2000). Such behaviour may include one or more of the following:

- **User modelling:** techniques which allow the interface agent to maintain a user profile which may include a knowledge about the users and their preferences.
- **User adaptivety:** techniques that allow the interface agent to adapt its behaviour to different users and usage situations.
- **The use of natural languages:** techniques that allow the interface agent to interact with their users in natural languages.
- **Dialogue modelling:** techniques that allow the interface agent system to maintain a dialogue with the user via natural language.
- **Explanation generation:** techniques that allow the interface agent to explain and demonstrate its suggestions to the users. This is important in constructing an appropriate and efficient justification for the interface agent's results and it requires a good knowledge of the way the user accepts such messages.

All the above techniques are useful in building an intelligent interface agent, and they depend on each other. For instance, if the interface agent has the ability to maintain a dialogue it should be able to interpret natural language utterances, and so on. In order to increase the intelligence of the agent, it must have the ability to learn. The agent can learn using one or more from the following four methods presented by (Maes 1994):

1. Looking over the user's shoulder. In other words, the agent can monitor the user behaviour when performing actions.
2. Utilising either direct or indirect user feedback. Indirect feedback happens when the user neglects the agent's recommendations and takes different action instead.
3. Giving the agent an explicit example to follow.
4. Acquiring advice from other agents in the system that assist other users with the same task.

Another learning alternative is to provide the agent with all the inputs and combine them with the desired output. The intelligence, learning techniques and their performance will be discussed in Sections 4.8 and 4.9.

4.6 Agent Architecture

As mentioned previously, the characteristics of agents differ, and affect the infrastructure of agents' architecture. There is no one standard design used for all agents. The architecture varies, depending on the nature of the task it is responsible for (Huhns and Singh 1997; Huhns and Singh 1998). For example, the architecture of a mobile agent differs from that of collaborative agents. As mentioned in Section 4.3, mobile agents migrate among networks; and operate on several kinds of machine across the Internet (Knapik and Johnson 1998). *Tutoring* collaborative agents (Rich, Lesh et al. 2002) work within a single server or network; the agent does not need to migrate or execute in other environment, thus the collaborative agent's architecture differs from that of the mobile agent.

Generally, there is no common architecture that is compatible with all types of software agents and which works in all circumstances (Huhns and Singh 1997; Huhns and Singh 1998). However, most of the agent architectures can be divided into four basic functions which are observation, recognition, planning and executing appropriate actions (Millier 1997). These functions can be performed by several of the software components such as:

- **The interface:** which acts as a mediator between the agent and its owner or user, or between the agent and other agents in the system in a multi-agent model (Rus, Gray et al. 1998). In both contexts the interface allows the agent to communicate with the outside world (Rus, Gray et al. 1998; Amant and Zettlemoyer 2000).
- **The controller:** which manages the flow of the operations within the agent. The controller is the part of the agent that decides the appropriate actions applicable to the current situation (Maes and Wexelblat 1996).
- **The sensors:** observe either the user's behaviour or the changes in the agent's environment (Maes and Wexelblat 1996; Millier 1997; Amant and Zettlemoyer 2000).
- **The effectors:** are to manipulate the observed data from the sensors (Maes and Wexelblat 1996; Amant and Zettlemoyer 2000). The effectors process the data sensed, and then prepare the agent to accept new data (Amant and Zettlemoyer 2000). That means the effectors manipulation tunes the agent's action to the states observed.

The above are the most common components that might be implemented in any agent architecture. Some agent architectures have a memory component in which the agent can store users' profiles for future utilisation. For instance, the *Apt Decision* agent presented by (Shearin and Lieberman 2001) intelligently profiles the users' preferences and saves them for future retrievals.

4.7 Agent life cycle

The agent is a computational system which is based on three major criteria, the agent's goals, the different states that need to be processed and appropriate actions produced in response to the states observed (Maes and Wexelblat 1996). These

criteria are necessary for the agent's life; the absence of one of them can affect the existence of the agent (Yap and keong, 1999). The agent goes through different phases during its life time; creation, execution, decision making and learning.

Firstly, in the agent creation phase, the agent is constructed with certain goals in mind (Franklin and Graesser, 1996; Maes and Wexelblat, 1996; Yap and keong, 1999). The agent's goals are presented the problems being solved (Maes and Wexelblat, 1996).

The second phase is the agent execution phase, which involves sensing and processing or manipulating the events targeted. The sensor is the most important part of the agent in that it is essential as a problem processor starter and as the agent's environment recogniser (Maes and Wexelblat, 1996). The sensors collect essential information about the environment that need to be manipulated (Maes and Wexelblat, 1996). In the agent's decision making phase, it decides appropriate actions taking into account the data sensed in the previous phase. Finally, the agent's learning phase could be considered the most important of all since it determines what to sense in the future. Additionally, the agent learns how to deal with similar situations from previous experiences and adjust its actions to get maximum reward.

The most important aspect of the fourth phase is that the agent learns and explores new ways to achieve its goals but does not change the goals it was built for. In other words, the agent has several tasks that it performs to accomplish its goals, and the agent's learning affects the performance of its tasks. Agents' sensing, learning and performance reflect the degree of intelligence that distinguishes them from the conventional programs (Maes and Wexelblat, 1996).

The agent life cycle will operate until one of the following happens (Yap and keong, 1999):

1. The agent stops accepting new tasks because drastic changes have been sensed in the environment.
2. The agent's owner terminates the agent's execution.
3. The agent's goals are no longer relevant to the users' tasks.
4. Higher priority agents, or newer versions of the agents replace the old agents in a system.

It appears that agents, like humans, have a limited life acting in the system (Yap and keong, 1999). However, this life can be extended by adding more intelligence to the agent and the discussion in the next section will cover different intelligent techniques.

4.8 Agent intelligence

The intelligence of the agent is the ability of an agent to adjust its behaviour according to the user's needs and to changes in the environment (Knapik and Johnson 1998). The intelligent nature of an agent can be considered as one of the properties that distinguishes it from other entities (Maes and Wexelblat 1996). As(Knapik and Johnson 1998) put it, the agent's intelligence is represented in:

- Reasoning about the world or domain in which the agent finds itself.

- Reasoning under uncertainty, with imprecise or incomplete data.
- Discerning patterns, learning from and generalising a specific aspect.
- Evolving to work appropriately with the situation observed.
- Adapting the structure of the agent to maximise its autonomy.

Several methods have been used to integrate intelligence within an agent; each method is specialised to handle one or more of the attributes mentioned above (Knapik and Johnson 1998). Three basic techniques have been used to develop the agent intelligence; fuzzy logic, neural networks, and genetic algorithms (Knapik and Johnson 1998).

4.8.1 Fuzzy Logic

Generally fuzzy logic based systems are computational systems that enable computers to deal with uncertain and imprecise data from a real world(Knapik and Johnson 1998; Mukaidono 2001). The fuzzy logic converging technique is used with the agent technology to enhance the agent's ability to cope with uncertain and new domains of data collected from either the users or the environment (Knapik and Johnson 1998). It can process ambiguous information based on Zadeh fuzzy set theory proposal (Zadeh 1994), which is founded on computational linguistics variables (Harris, Moore et al. 1993). Fuzzyfying the crisp or numerical variables is considered the first component of the fuzzy logic controller (Harris, Moore et al. 1993). The second part of the fuzzy system is the inference, which contains fuzzy rules that are responsible for producing appropriate fuzzy output. Defuzzification of the fuzzy output is the third component of the fuzzy logic model, and it changes the output to crisp values (Harris, Moore et al. 1993). Figure 4.1 shows the main components of the fuzzy logic system.

The linguistics variables used in fuzzy logic are the words used to describe or criticise a specific entity (Knapik and Johnson 1998). For example, the linguistic variables that might be possible to characterise distance could be of the form of Very Close, Less Close, Close, Far, Really Far and etc. (Harris, Moore et al. 1993). Fuzzy logic linguistics variables differentiate it from the other techniques and make it suitable for several applications mentioned by different authors (Knapik and Johnson 1998; Mukaidono 2001):

- Experts systems.
- The production, processor of controlling systems such as controlling the system temperature.
- Assessing in market prediction and trading.
- Inventory control.
- Route planning.
- Patterns detection.
- Controlling for robots and Roberts arms.

Expert systems are used for problems diagnosis, planning, prediction, natural languages and user interface (Knapik and Johnson 1998). Expert systems applications

are considered to be the most significant beneficiaries of the fuzzy logic (Knapik and Johnson 1998).

Decisions in fuzzy logic systems are based on the identification of inputs in the form of linguistic variables driven from the membership functions (Lin 1994). The membership functions are formulas used to determine the degree of the relationship between the input variables (crisp) and the fuzzy input sets (Lin 1994). The fuzzyfication process consists of calculating the membership function of the crisp input variable that produces values between (0- 1). Different membership functions have been established to support a mapping of crisp input variables to their correspondence (0-1), such as triangular, trapezoidal, quadratic, gaussian and crisp or boolean (Harris, Moore et al. 1993). The selection of the function depends on aspects of the variables, for instance, whether the crisp values are infinite. Also, membership function or the fuzzy sets selection affect the output, since the fuzzy output is directly dependent upon the shape of the function selected (Harris, Moore et al. 1993).

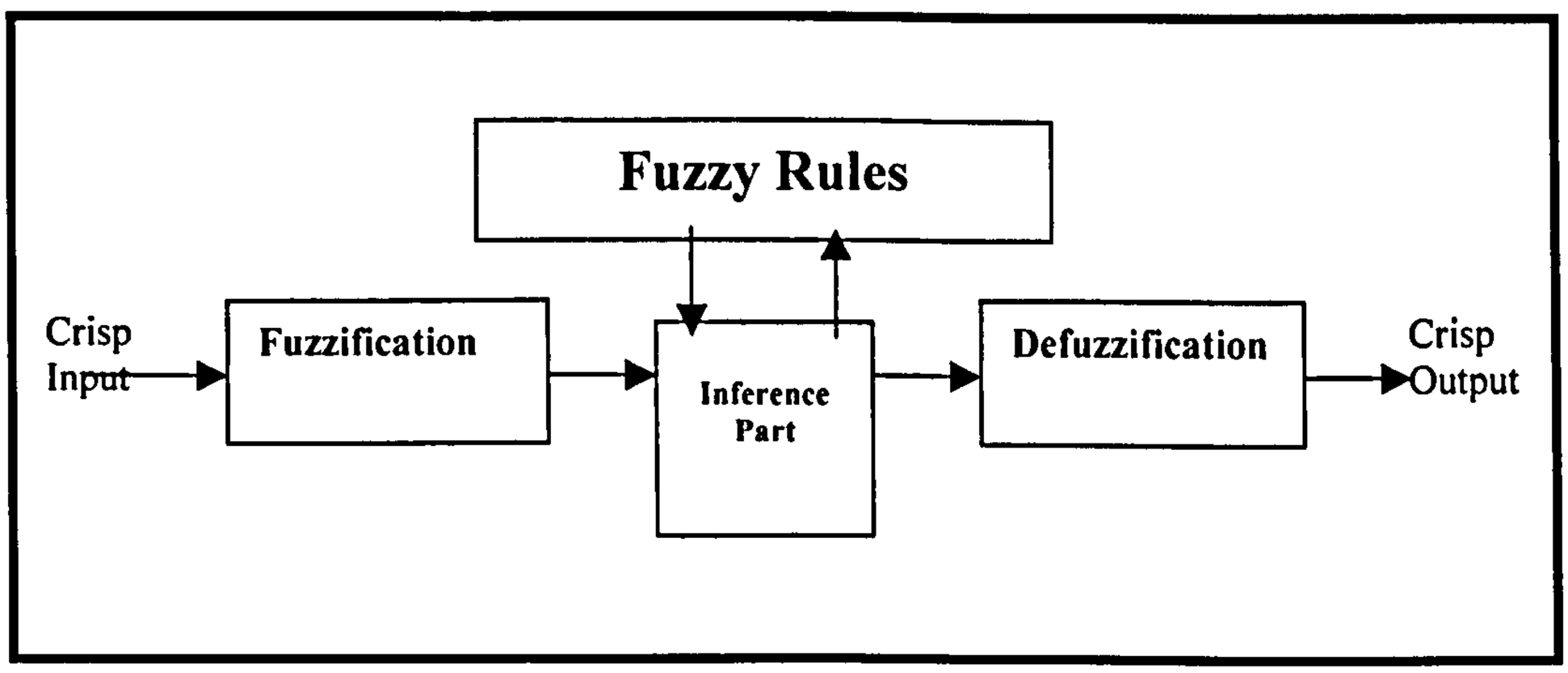


Figure 4.1: Generic Fuzzy logic Controller

The second part of the fuzzy logic system is the fuzzy inference which contains the fuzzy rules. The fuzzy rules are conditional linguistics statements or equivalently fuzzy relationships between the input and output fuzzy sets (Harris, Moore et al. 1993). Each rule can be formed in the IF-THEN logic format and matches the input variable with the appropriate output (Lin 1994). All the logical operations can be applied to form the conditions of a fuzzy rule, operations such as logical AND, OR, difference. One or more fuzzy rules can participate in generating the fuzzy output.

Different aggregation functions, such as centroid (Lin 1994), min and max (Harris, Moore et al. 1993) might be applied to the output calculated from the fuzzy inference part in order to come up with a specific value.

In order to describe how the fuzzy logic is processed, we will consider the fuzzy colouring model established as part of this research by (Abulkhair and North 2003). Abulkhair and North (2003) fuzzy colouring model was used to measure the colour contrast between the text and background colour in a web pages. The fuzzy colouring model involves two phases, the colour intensity measurement and then the colour contrast measurement to produce a reasonable usability evaluation. In the intensity

measuring phase, the text or background colour is input. The input is fuzzyfied to an equivalent fuzzy linguistic variables form using a triangular membership function (see Chapter 6, Section 6.5.2). For this fuzzy model the linguistic variable's range is: *Dark, MidDark, Medium, MidLight* and *Light*. There are around 125 fuzzy rules to derive the colour intensity measure, and these can infer the intensity of the colour even if its intensity lies between two consecutive linguistic input variables. The colour intensity inference depends on applying all the rules that involve the colour intensity fuzzy input.

For instance, if the text colour is (#00 33 ff) hexadecimal, then in the RGB colour analysis: red is 00, green is 51 and blue is 255. Since 33 hex and 51 decimal lie in the overlap area as Figure 4.2 shows, two fuzzy rules are selected, those which contain *MildDark* and *Dark* for the green part of the RGB. Therefore, two membership functions are calculated and two fuzzy rules are involved in the output inference, in order to produce a reasonable solution, the first fuzzy rule is:

IF (Red IS Dark) AND (Green IS Dark) AND (Blue IS Light) THEN (ColourIntensity IS Darkish)

Whereas the second one is:

IF (Red IS Dark) AND (Green IS MidDark) AND (Blue IS Light) THEN (ColourIntensity IS Medium)

Figure 4.3 illustrates the range of the colour intensity fuzzy output. In order to come up with a crisp value the fuzzy output is aggregated by using a centroid function and then defuzzified. The crisp colour intensity value defuzzified is equal to 47.5. The details of this fuzzy model can be seen in Appendix *C*.

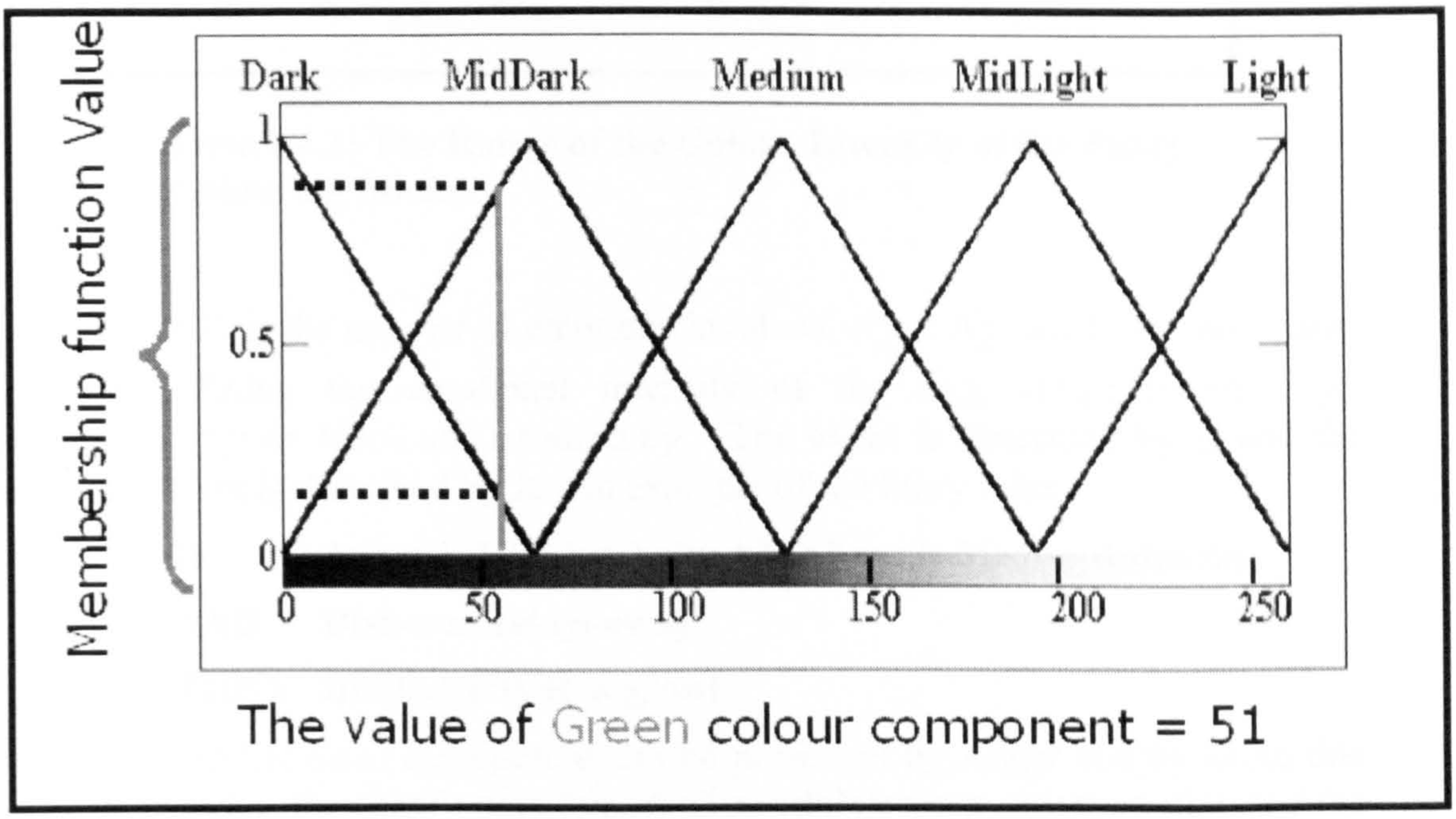


Figure 4.2: Two Membership Function Values for Each Number Lies in the Overlap Area

The Fuzzy Logic Adaptive Model of Emotion (FLAME) is an adaptive agent that acts as a pet (El-Nasr, Yen et al. 2000). The (El-Nasr, Yen et al. 2000)'s agent focuses on the emotions produced by the dog and it is affected by the users' responses. The agent fuzzy rules are of the form:

IF **emotion$_1$ is A$_1$ AND**

emotion$_2$ is A$_2$ AND ..... AND emotion$_k$ is A$_k$

AND Event is E

AND Cause (E,B)

THEN **BEHAVIOUR is F**

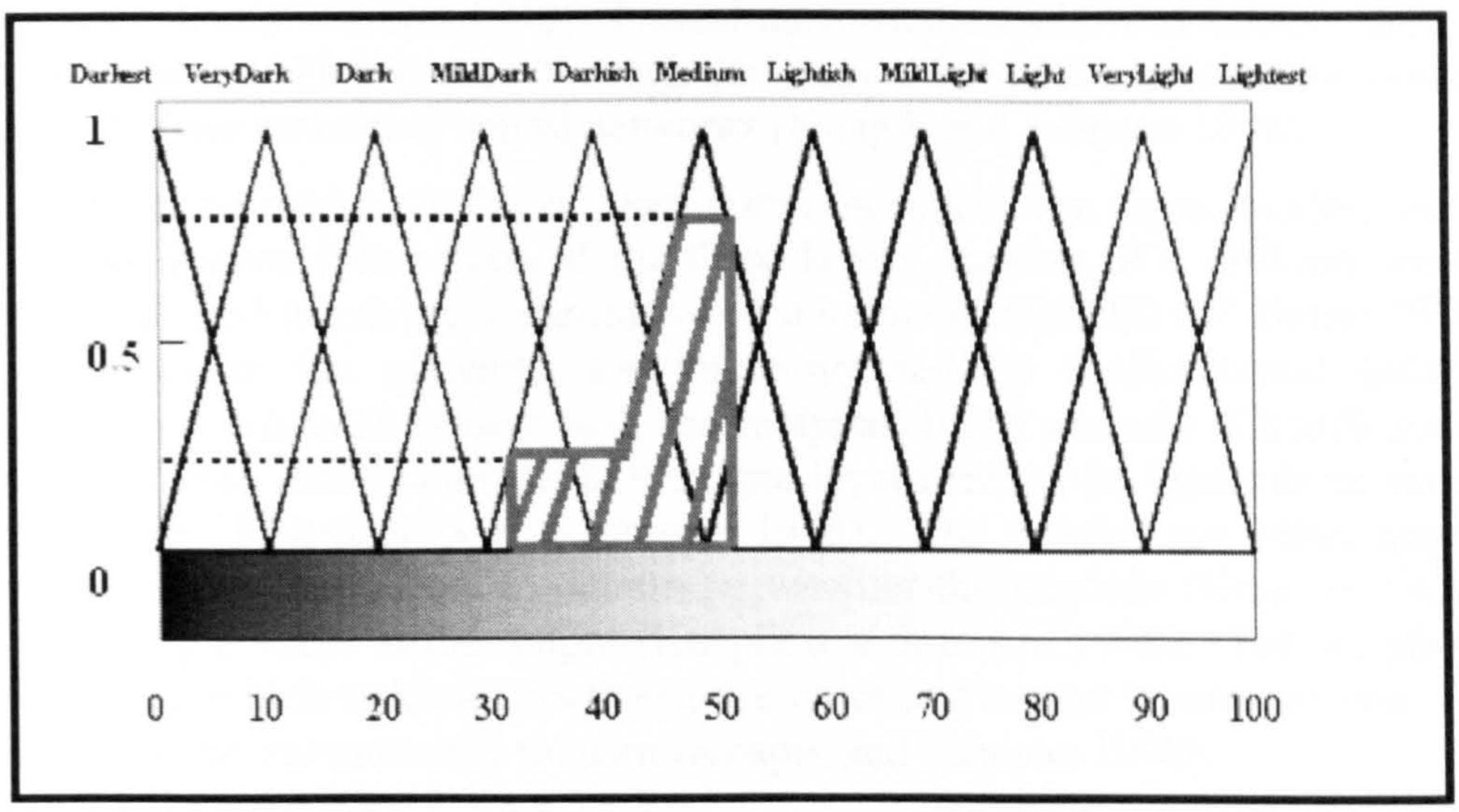


Figure 4.3: The Range of the Colour Intensity of the Fuzzy Colouring Model

Where K is the number of emotions involved. A_1, A_2 and A_k are fuzzy input sets defining the emotional intensity of the dog, *HighIntensity*, *LowIntensity* or *MediumIntensity*. The event is described by E and the cause of the event is described by B. An example of the fuzzy rules

IF **Anger is HighIntensity AND Fear is MediumIntensity**

AND **Dish-was-taken-away**

THEN **BEHAVIOUR is growl**

The *HighIntensity* emotion can be presented by Anger and by using this form of fuzzy rules, the agent's negative emotion inhibits the positive emotion and the behaviour is affected. This example shows how the fuzzy logic can be used to control the behaviour of the adaptive agent.

As discussed earlier, fuzzy logic has been used in experts system for users' interface problems (Knapik and Johnson 1998; Abulkhair and North 2003).

Abulkhair and North's study (2003) showed that fuzzy logic works effectively in measuring the colour intensity and then producing a proper usability rate. In order to manage a wide and imprecise range of web usability guidelines, it is reasonable to use fuzzy logic to control the evaluation and to involve a large number of current usability guidelines. That is why fuzzy logic was used to control the usability measurement.

4.8.2 Neural Networks

The neural networks is modelled on neurons in the brain (Hertz, Krogh et al. 1991). Neural networks mimics certain aspects of the architecture of a human brain and applies the brain's biological mechanism in learning (Knapik and Johnson 1998). This imitation enables the neural network to learn how to do specific tasks rather than codifying every single step in the algorithm that executes the task (Caudill and Butler 1990; Knapik and Johnson 1998). Neural networks learn how to achieve their task or tasks by training. This is done through the recognition of the information patterns, an important characteristic of neural networks (Knapik and Johnson 1998).

Neural networks consist of three common layers, the input, hidden and output layers (see Figure 4.4). Each of the three layers consists of a different number of elements (nodes) which are sometimes called neurodes (Caudill and Butler 1990). All the neurodes in the networks are interconnected by unidirectional links called synapses that link each output with the recipient input neurode (Caudill and Butler 1990; Lin 1994). Each neurodes in the input layer receives its inputs from sources via weights (Lin 1994; Knapik and Johnson 1998). The weights are either negative or positive numbers calculated to determine whether the neurode "fires" or "succeeds" by producing a value at the output (Knapik and Johnson 1998). The weights of the hidden layer, which is located between the input and output layers, are crucial to the ability of the neural networks to learn (Knapik and Johnson 1998).

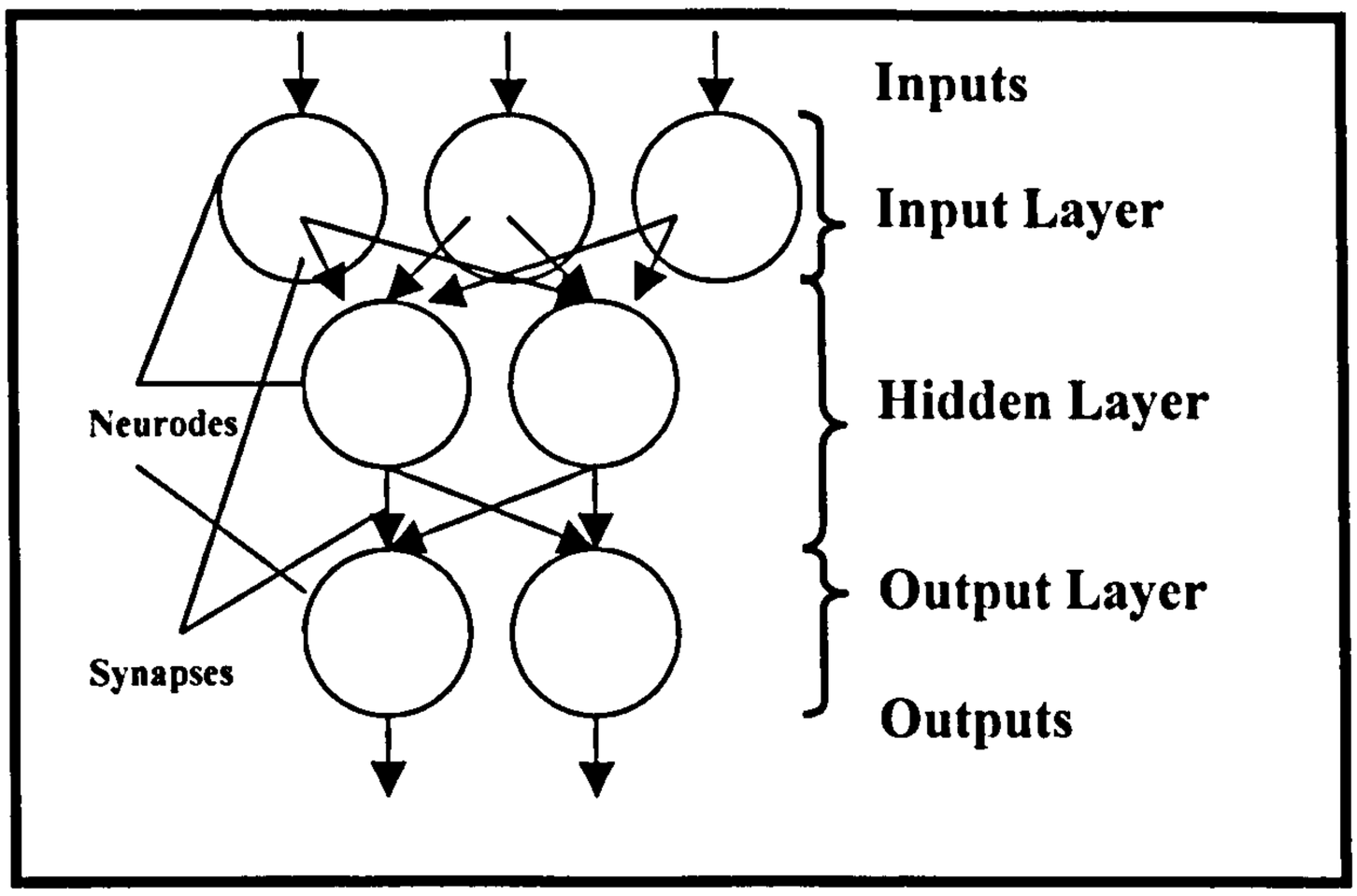


Figure 4.4: Generic Neural Networks Architecture

The neural network is modelled on three basic entities, the model of the neurodes, the synaptic interconnections and structures, and the training rules (Lisboa 1992; Lin 1994). Knapik and Johnson (1998) summarise neural networks training; the weights are initialised to some values and the output results are compared to the output desired. The difference between them is the error. This process should be repeated until the output goals are reached.

Pattern recognition, transformation and classification are the most common neural networks applications (Rich and Knight 1991; Lisboa 1992). Therefore, it is reasonable to utilise neural networks in applications such as sales forecasting, defence systems for identifying the friends from the enemy, optical character and hand writing recognitions (Knapik and Johnson 1998) and other examples cited in (Lisboa 1992). Neural networks store data in weight patterns (Caudill and Butler 1990) which are still at too primitive a stage to be applied to dynamic inferencing (Rich and Knight 1991). Thus, this technique for producing intelligence was not utilised in this research, because the dynamic nature of web page designs and their measurement mean it is inappropriate to do so. In addition, the research aims to customise the usability evaluation to the webmaster's feedback, which is dynamic too.

4.8.3 Genetic Algorithms

A genetic algorithm is a computational model with a basis in the biological process of evolution (Forrest 1996; Melanie 1996; Knapik and Johnson 1998). Genetic algorithms are simply stored in a computer's memory maintaining a gene pool for individual parameter settings and performance metrics or goals. The creation of new components is made by producing more copies of the successful components that met the goals, and deleting the less successful ones. The copies are produced by two operations; mutation and crossover (Forrest 1996). A fitness function is utilised to optimise the parameters produced to best suit the agent goals (Knapik and Johnson 1998). The generic algorithm performance has three components mentioned by (Forrest 1996):

- Large population initialised randomly to provide independent sampling.
- High-fitness individuals are preserved through selection.
- Crossover combines partial solutions from different strings onto the same string to maintain a population of candidate solutions.

Genetic algorithms are highly recommended for optimisation tasks such as searching in a wide sample of data, scheduling problems (Zheng and Kiyooka 1999) and sorting networks (Forrest 1996). It is not appropriate to use this type of algorithm in this research. The multilingual automated web usability evaluation agent implemented needs extra information such as webmaster's feedback to customise its assessments. On the other hand genetic algorithms do not use extra information, and as a result have a slow convergence rate (Zheng and Kiyooka 1999).

4.9 Learning Techniques

A machine cannot be called intelligent unless it is able to learn to process the problem it is solving in different ways, or to adapt to new situations (Rich and Knight 1991). However, (Anthony and Bartlett 1999) defined learning as the adjustment of the system in response to data generated by the environment. Adaptive ability is one of the important agent's characteristics, as it involves the agent's ability to learn from the user or from the environment. The learning techniques are classified into three major groups; graded, supervised and unsupervised learning (Caudill and Butler 1990).

Supervised learning provides the learning algorithm with the inputs and corresponding desirable outputs (Caudill and Butler 1990; Hertz, Krogh et al. 1991; Lin 1994; Knapik and Johnson 1998; Serban 2003), it is like the tutoring or teaching in schools. Unsupervised learning is a self organising learning technique that might be given an input but the output targeted is not provided (Rich and Knight 1991; Knapik and Johnson 1998). It works efficiently with redundant input data, which gives the algorithm the opportunity to explore the input pattern (Hertz, Krogh et al. 1991). Graded learning is not provided with actual output pattern but only given a feedback to indicate its performance (Caudill and Butler 1990).

These algorithms can be presented in different forms such as backpropagation networks, Boltzmann machine, Hopfield networks, Hybird learning algorithm, On-line supervised learning, Reinforcement Learning, etc. and they have been discussed by different authors (Caudill and Butler 1990; Hertz, Krogh et al. 1991; Rich and Knight 1991; Lisboa 1992; Harris, Moore et al. 1993; Lin 1994).

4.9.1 Reinforcement Learning

Reinforcement learning is the way of exploring the environment and improving the behaviour of the agent by giving it feedback about his performance (Serban 2003). But even though reinforcement learning has been considered as a form of supervised learning (Hertz, Krogh et al. 1991; Serban 2003); (Blumberg, Downie et al. 2002) classify it as an unsupervised learning. Neither of these considerations is accurate, since reinforcement learning is aimed at maximising the reward in the absence of a teacher (Blumberg, Downie et al. 2002), and that contradicts the first suggestion. It is not unsupervised learning because to perform its goals reinforcement learning should be provided with a reward, which is one of the four common reinforcement learning elements (Sutton and Barto 1998). However, it can be classified as graded learning since it just needs a feedback from the environment instead of being provided with the input and associated target output values. A differentiation between the reinforcement and supervised learning has been expressed by (Lin 1994). The feedback is for evaluating the output but not for instructing the system using the reinforcement learning (Hertz, Krogh et al. 1991). The feedback could be in the form of *yes/no* (Hertz, Krogh et al. 1991), or by either positive or negative real values (Rich and Knight 1991).

Reinforcement learning is founded on four main elements; the policy, reward function, value function and optionally the model of the environment (Sutton and

Barto 1998). Policy defines the learning agent's behaviour at a given time, and the reward function is used to determine the goals of the reinforcement learning. Reinforcement learning has been utilised by different authors (El-Nasr, Yen et al. 2000; Seo and Zhang 2000; Seo and Zhang 2000; Isbell, Shelton et al. 2001; Shapiro, Langley et al. 2001; Blumberg, Downie et al. 2002) to customise the agent's behaviour. Figure 4.5 depicts the reinforcement learning procedure containing the steps presented in (Isbell, Shelton et al. 2001).

The Q-learning algorithm is a specific type of the reinforcement learning that can map a state to an action to maximise the reward (Sutton and Barto 1998; Isbell, Shelton et al. 2001). Reinforcement learning has been used to learn user's preferences by observing the user's reactions to the filtered documents and to adapt the most significant terms that best represent user's interests (Seo and Zhang 2000) while *Icarus* is an agent architecture that embeds a hierarchy within the language for specifying agent's behaviour (Shapiro, Langley et al. 2001). Different authors (Kauchak; Brown, Jr. et al. 1998; Kauchak 2001; Blumberg, Downie et al. 2002; Tumer, Agogino et al. 2002) utilised reinforcement learning to customise the behaviour of their agents in different ways. Thus, it is reasonable to utilise Q-learning in this research, since it is widely used in personalisation and shows effective results in customisation. Because the multilingual automated web usability evaluation agent developed in this study needs to customise its evaluation to the webmaster's feedback, it needs a learning technique that can provide a certain degree of personalisebility which can be achieved by using Q-learning.

1. The agent senses the state of the environment and executes the appropriate action.
2. The agent receives a reward for its performance.
3. The agent chooses actions to maximise the reward, and maps from states to actions that achieves the agent's goal.

Figure 4.5: Reinforcement Learning Generic Procedure

4.10 Conclusion

It is clear that agent technology can act on behalf of users and can support them to resolve a set of tasks that might tedious, repetitive, dull or hard to perform. Agent technology can be utilised in various applications, but this technique should have a

certain degree of intelligence in order to work efficiently. There are several techniques to implement the intelligence of the agent such as fuzzy logic, neural networks and genetic algorithm. Additionally, many authors have suggested developing intelligent agents with the combination of one or more of the techniques mentioned earlier. These techniques also might be supported with several learning approaches to increase the learnability and the adaptivety of the agent developed.

Integrating agent technology in the automated web usability evaluation is considered a crucial step for developing intelligent automated web evaluation. The current study will utilise the agent technology to analyse, evaluate and customise the agent's evaluation to the webmaster's preferences. The analysis involves the identification of web metrics under scrutiny. However, the evaluation will be based on the current usability guidelines in combination with user satisfaction, which is part of this research and its methodology will be discussed in the next chapter. Customising the evaluation to the webmaster's feedback involves the utilisation of three learning approaches which will be discussed in Chapter 6.

Chapter Five

The User Satisfaction Methodology

5.1 Introduction

The existence of several web usability evaluation methods, discussed in Chapter 3, increases experimenters' capabilities in discovering interface problems. Using more than one usability evaluation method is a valuable technique to produce a reliable assessment. Each method tries to handle some aspect of usability which the other methods cannot. Among these, user testing and user satisfaction is a common usability testing technique which involves investigating web interface features that have a direct impact on users.

In order to explore the users' preferences, user satisfaction was selected because it is considered to be the closest technique to the users; it can be applied to discover users' favourites, thoughts and working habits (Nielsen, 1993; Ivory, 2001; Ivory and Hearst, 2001). Furthermore, it is used widely to identify differences in users' culture (Yeo, 1998; Lee, 1999; Marcus and Gould, 2000; Goonetilleke et al. 2001; Marcus, 2001). A remote user satisfaction technique in particular was chosen because it allows users to be tested in their normal environment rather than in a laboratory.

This chapter discusses remote usability testing and how it was used to discover both Arabic and English users' preferences as regards some of the main features of the web page layout. To enhance the information presented within the web page, cultural needs must be reflected in the web page design. Due to the lack of any usability guidelines especially intended for Arabic web pages, with their different script and writing direction, it was necessary to carry out a user satisfaction to acquire knowledge of Arabic users' likes and dislikes. Additionally, it is appropriate to have similar knowledge about English users' preferences to validate the current usability guidelines.

Remote usability testing was used, since it is appropriate for web environment, because it is reasonable to test in working, studying and home environments. In other words, the test does not have to be run in the laboratory to get effective results; instead, the participants' interaction with the web should be measured in the participants' usual working environment. Changing the interface environment influences the design and the users' choices, since the users' environment is considered one of the interface aspects that might affect the interface design (see Chapter 3, Section 3.4). One of the main aims of this research is to recognise the users' preferences in any environment, home, work and study places etc. to produce a general purpose users' preferences. The details of the remote user satisfaction are discussed in this chapter.

5.2 Stages for Conducting the Remote Usability Testing

As mentioned in Chapter 3, Section 3.6.2.1, usability testing in general passes through several stages before getting the results. However, the stages for this remote usability testing are as follow:

5.2.1 Test Plan:

Five major goals were set for this remote usability testing; attractiveness, scanability, legibility, readability and different aspects' utilisation. These goals were set to test both Arabic and English users in their own language, since localising the usability testing language representation is considered to be of the factors, which affect the usability testing success (Yeo, 1998). In order to shorten the test, the remote user satisfaction was performed in three individual phases, each phase has its own task and all of them work cooperatively to achieve the usability testing final goals. The content specification of all the three stages of the usability testing is discussed in relation to the questionnaire specification from this chapter. However, the layout of the web pages that contained the test were considered with reference to the current guidelines in web presentation, such as selecting dark text over light background colour (Shneiderman, 1998; Nielsen, 2000; Lynch and Horton, 2001; Nielsen and Tahir, 2002).

5.2.2 Selecting Participants:

The remote usability testing involved in this study aimed to identify general usability principles rather than specified usability guidelines. The participants targeted were habitual web users, since the need is to learn about general users' preferences. For this usability testing, there is no need for demographic information about the participants except their native language, since the cultural differences affect the users' preferences, thoughts and feelings (Lee, 1999; Marcus and Gould, 2000; Marcus, 2001; Badre, 2002). However, children were not enrolled as part of the participants, since they have different concerns and special web page presentation.

5.2.3 Preparation:

Six questionnaires were created, three for each language, and classified to cover different legibility and other relevant aspects. The questionnaires were divided into colour, text, title and alignments. The English questionnaires were built up first, and revised several times by a native English speaker. Then same the questionnaires were translated to Arabic by a native Arabic speaker who is fluent in English. The questionnaires' web pages were constructed in a form of multiple choices, yes/no questions and dropdown menus, depending on the question type. Several questions' forms are shown in Figure 5.1 and Appendix *D* has all the Arabic and English questionnaires. For instance, a multiple choice question is used for question 4,

whereas question 5 represents the yes/no type and a drop menu is provided for question 6.

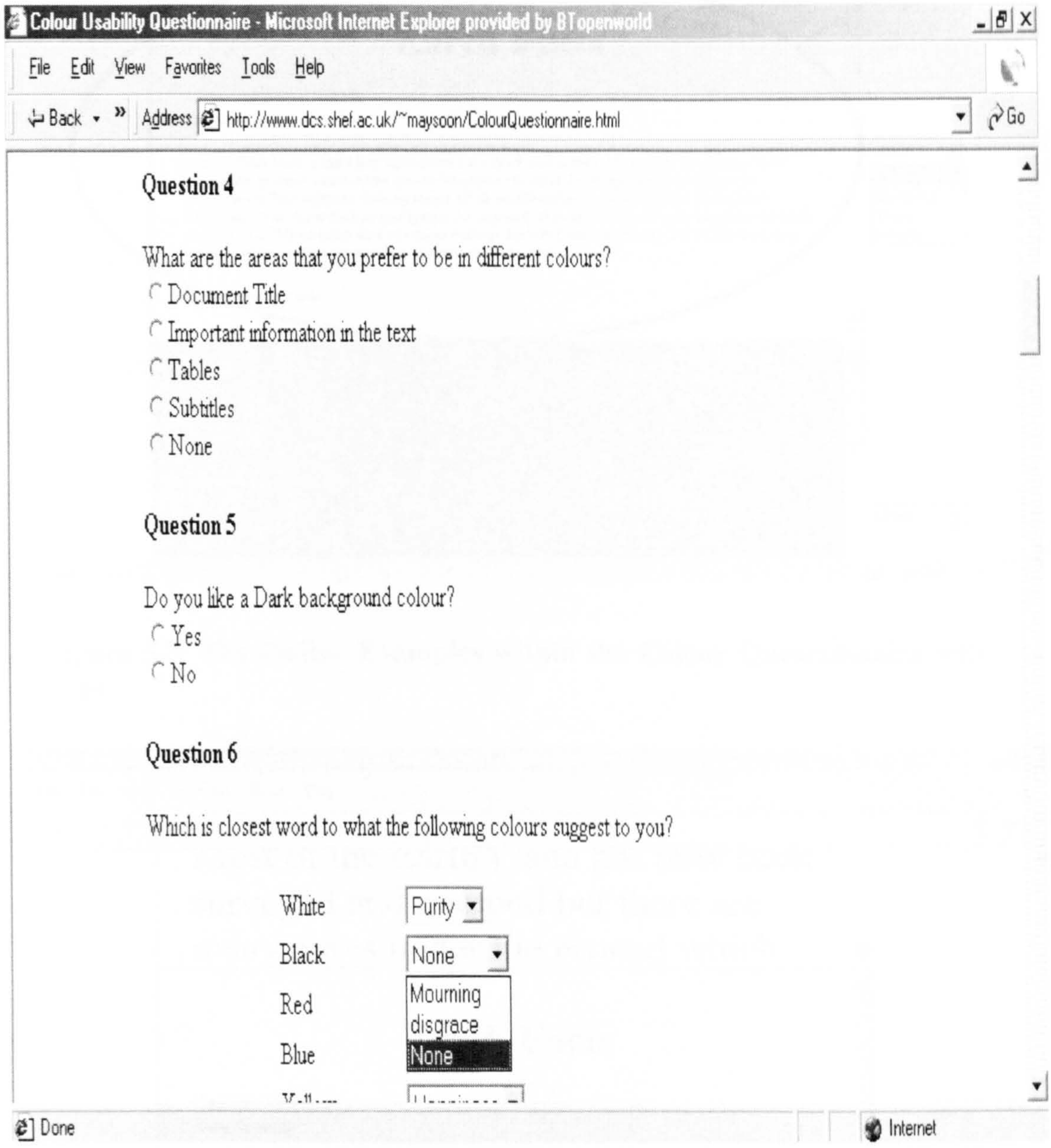


Figure 5.1: The Forms of the Questionnaire's Questions

Also, usability testing involved the development of online examples to make it easier for the participants to choose their favourites. Figure 5.2 shows how these examples were presented. At the end of each questionnaire there is an appreciation message which thanks the users for their participation. After the participants completed the questionnaire, they can submit it by pressing the submit button which appears at the bottom of the Figure 5.3.

After completion of the questionnaires web pages development, the participants were sent an email which politely encouraged them to participate in building the users' preferences and contained the URL user satisfaction address.

Figure 5.2: The Online Examples within the Colour Questionnaire web Page

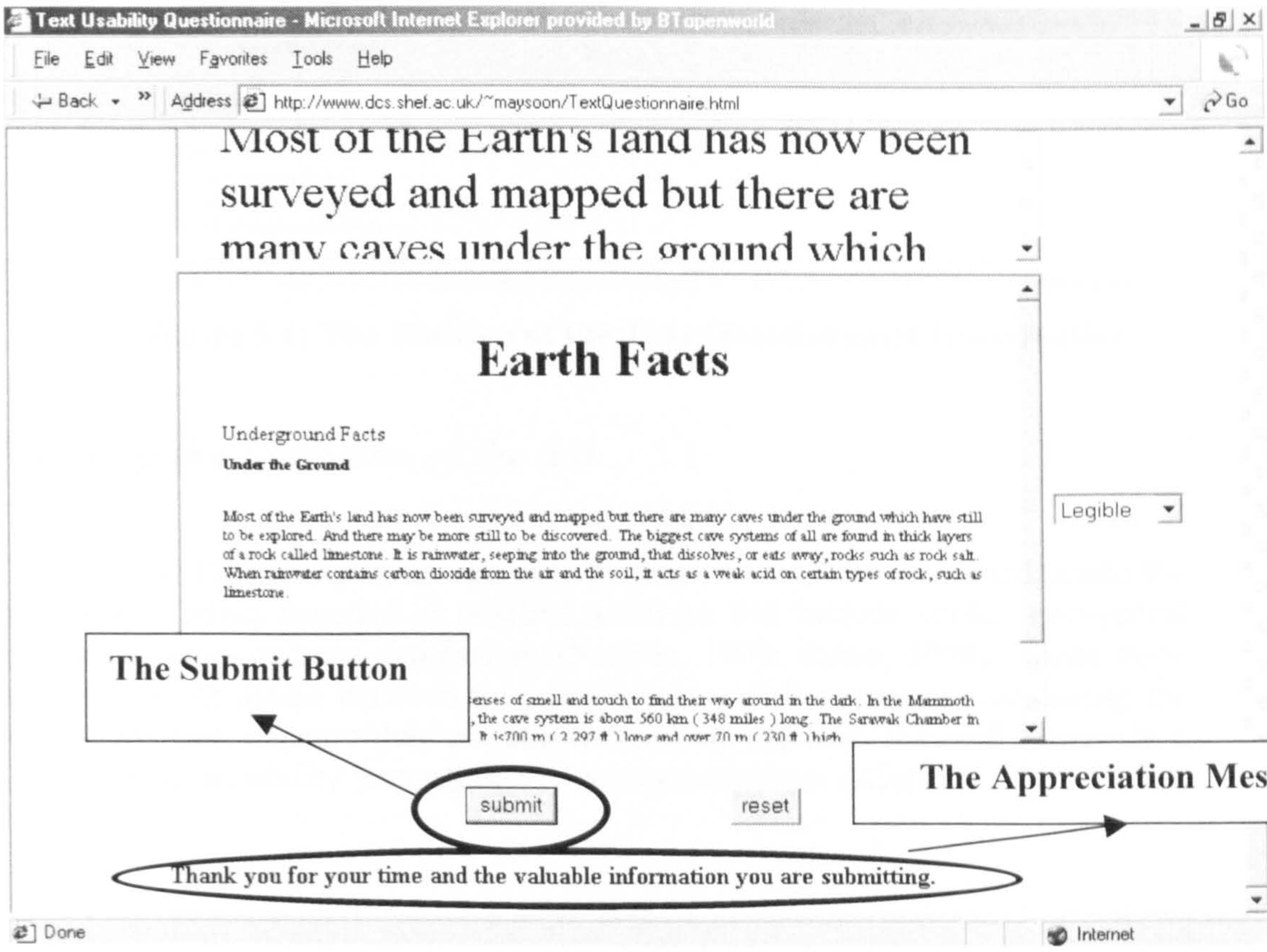


Figure 5.3: The Representation of the Submit and Appreciation Message

5.2.4 Introduction:

An introduction is located on the top of each questionnaire's web page, six introductions were written, three for each language, Arabic and English. Each introduction gives the participants a general idea about the colour, text and title and alignments effects on the web design and how these different issues might affect user acceptance. The general idea makes the users concentrate on the main topic of the running user satisfaction, which is reflected in their answers. Figure 5.4 illustrates the position of the questionnaire introduction such as text questionnaire as presented in the Figure.

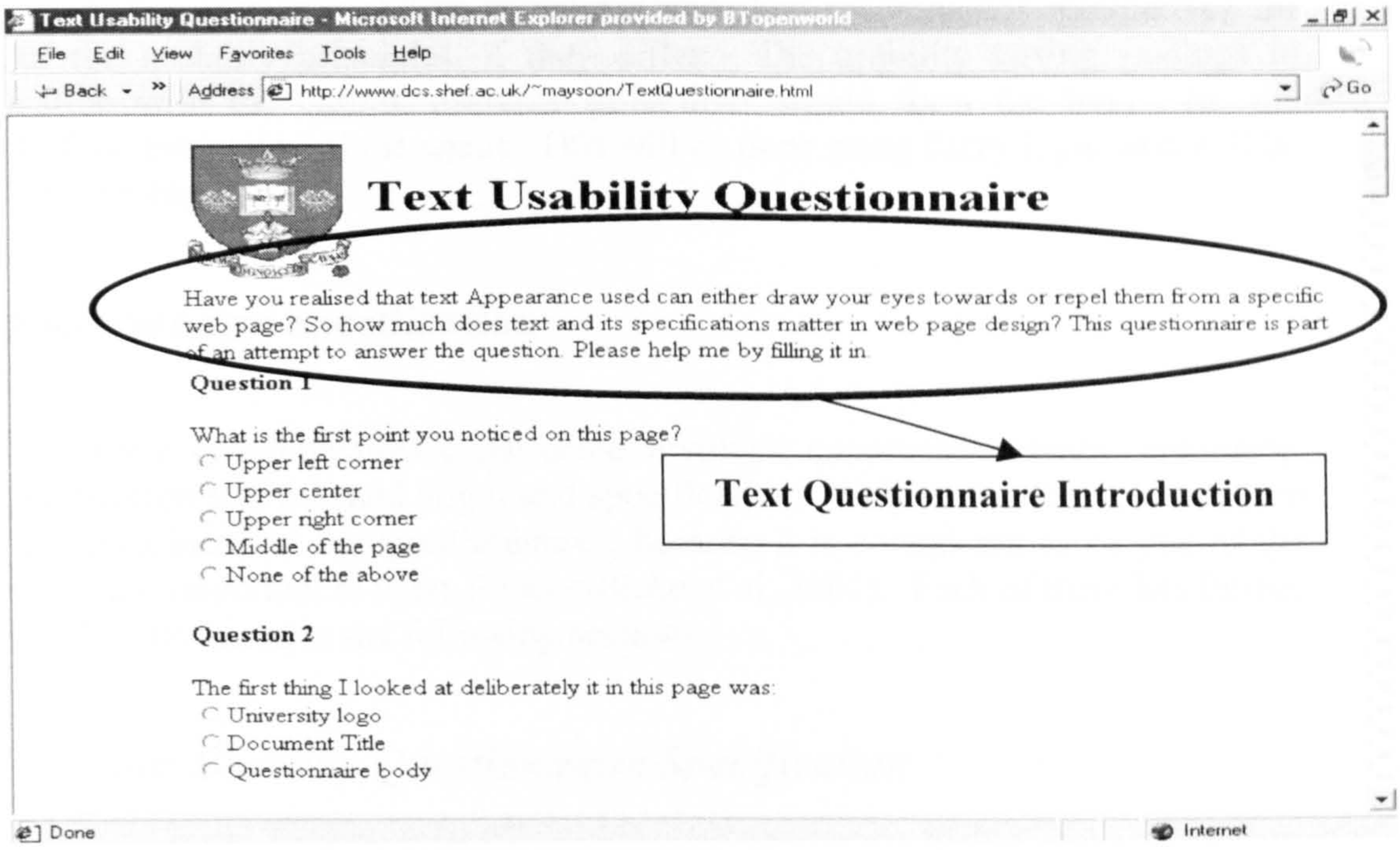


Figure 5.4: The Position of the Text Questionnaire Introduction

5.2.5 During the Execution of the test

To increase the reliability of the test, the experimenter must not influence the participants with either negative or positive reactions that include verbal, non-verbal signals, such as faces or hand expressions (Nielsen, 1993; Rubin, 1994). Since there is absolutely no interaction between the participants and the experimenter during the remote usability test, because they are not in the same location, this will guarantee a certain degree of reliability according to recommendations (Nielsen, 1993; Rubin, 1994).

5.2.6 Debriefing

None of the usability testing questions needs to be justified, so there are no comment spaces within the usability testing web pages. However, participants could send their comments if they have any, via email when the results of the user satisfaction are submitted. These comments do not affect the results of the user satisfaction evolved in this research, but they might be considered in future studies.

5.2.7 Transforming the Data into Findings

A comprehensive analysis will be covered in Chapter 7, which associates these results with the current usability guidelines and then recommends alternatives for some of the usability principles, if they differ. The usability testing findings in combination with the current usability guidelines would form the basics for the control of the evaluation of the agent. This will be done using fuzzy logic, and will be described in Chapter 6.

5.3 Questionnaire Specification

The three questionnaires constructed involved consecutive issues: colouring, text specification and title and alignment specification. Web readability is a common issue involved in all of the questionnaires, because it is considered to be one of the factors that are important to users (Goonetilleke et al. 2001). Each of them has further details and is discussed in the following sections.

5.3.1 Colour Usability Questionnaire Specification

This questionnaire is intended to study different web usability aspects related to colour and how they might affect the usability of any web page. All the questions are focused on three major issues; readability, legibility and attractiveness. In addition, there are some other questions that affect the readability such as the utilisation of the colours in different parts of the web page. The meaning of the colours used might vary from one culture to another (Russo and Boor, 1993; Marcus and Gould, 2000; Badre, 2002); consequently they can be used differently to display the information.

The meanings for each colour included in the questionnaire are derived from the most popular colour meanings in both Arabic and English. Nygaard and Fahrmeyer (2001) described red as one of the powerful colours that can be used for attention, while it can mean joy in Arabic culture. There are some feelings associated with particular colours such as: green is envy, and blue is sad as mentioned by (Terpstra and Sarathy, 1994), whereas spring is the other common indication for green and calm for the blue. Black is mostly associated with mourning in western countries (Terpstra and Sarathy, 1994; Nygaard and Fahrmeyer, 2001) whereas in the Arab culture it might indicate disgrace. Usually, white represents purity and cleanliness, but in the Eastern cultures white means death as stated by (Nygaard and Fahrmeyer, 2001).

Yellow symbolises the brightness of sunshine (Parzek, 1997; Nygaard and Fahrmeyer, 2001), and means happiness according to other authors (Parzek, 1997).

Individually, each colour has its own symbolic meaning; the combinations of different colours might change their meanings. Combinations might produce their own symbols, which can be added to the original colours' meanings. The consideration of colour combinations is very important in the web design because that might significantly affect the users' perception. As well as the symbolism, colour combinations can also affect the readability of the web pages. Therefore, there are several questions in the questionnaire concerned with web page legibility in relation to colour contrast. These questions involve measuring the contrast associated with the same four examples to test different aspects of the web design in order to have a consistent assessment from the users.

There are other questions that study the utilisation of colour in a web design. All the colour questionnaire questions are applied to both Arabic and English to study the web design differences which might affect the web usability. Since the presentation of any language is associated with a different culture and values, it is important to focus on these issues to produce colour usability guidelines relevant to a particular language.

5.3.2 Text Usability Questionnaire

This questionnaire starts with some questions about the ways users look at the web page. The questions focus on the first point the users' eyes scan, and so relate to the title and text body. The text specification is the focal point of this questionnaire. Text specification involves several issues dealing with the appearance of the text, in terms of letter size and shape. The text involves both subtitles and the body of the text.

This questionnaire ends with the question about the users' preferences in text presentation and headings based on the legibility and faster scanning. In addition to discovering the ways the users scan web pages, this is important in organising the information.

5.3.3 Title Usability Questionnaire

Many of the usability guidelines for web page's title are discussed in the literature, however most of these recommendations are for the browser titles but not for the document title. It is good to have a detailed specifications for the browser title; on the other hand there should be at least some description of the document title. Part of this questionnaire is aimed at building basic principles for the document title such as the length of title suitable for Arabic and English users.

This questionnaire will also include the alignment of the headings and the text, and what the users find reasonable for fast scanning and also reading. Users might lose concentration when reading if the web page is crowded with links; therefore, Nielsen (2000) recommended the inclusion of a small number of highly relevant links

in the web page. But do the users agree with this idea or do they like lots of links included in the web pages? This kind of query will be part of this questionnaire.

5.4 Remote Usability Satisfaction Reliability

In order to utilise remote usability testing findings in building the fuzzy rules, as it will be discussed in the next chapter, it is necessary to make sure that these findings are reliable. Two Cornbach's alpha reliability tests were used, since that is recommended by (Burns, 2000; Bryman, 2001; Pallant, 2001) to assess the internal consistency of the items and then ensure internal reliability. The first is driven from Cornbach alpha coefficient, which is reasonable at finding the correlation for more than ten items tested. This kind of test shows acceptable internal consistency, with a reported Cornbach alpha coefficient of 0.73 for the colour usability testing, which is an acceptable degree of consistency (Burns, 2000; Pallant, 2001). The second test performed is to report the mean inter-item correlation instead of alpha coefficient, because it is recommended for ten items and less (Pallant, 2001). The mean inter-item correlation for the text and title usability testing produced 0.3 and 0.4, which are both acceptable since they are between 0.2 and 0.4 as recommended by (Pallant, 2001). Table 5.1 shows the Cornbach alpha reliability scale for the text and title usability testing, which involves the Arabic and English usability testing.

User satisfaction	Alpha's Value
Text user satisfaction	0.3
Title user satisfaction	0.4

Table 5.1: Cornbach's Alpha Reliability Scale for Remote Usability Satisfaction

5.5 Remote Usability Satisfaction Validity

The validity is the degree in which the method measures what it is supposed to measure (Pallant, 2001). The remote usability testing method used in this research is designed to study the preference differences between Arabic and English users. Thus, it reflects the cultural differences. For instance, the usability testing method used provided descriptions of the characteristics of different web aspects such as attractiveness, scanability, legibility, readability and different web aspects usage such as colours. The reliability degree established in the previous section indicates that the remote user satisfaction used in this research is valid and its findings can be used in establishing the foundations for Arabic usability guidelines and validating the current usability guidelines for English.

The detailed description of the remote usability testing findings appears in Chapter 7. The next chapter contains a discussion of the second part of the methodology which describes the specification of the automated web usability evaluation applied using agent technology.

Chapter Six

Multilingual Interface Agent's Methodology

6.1 Introduction

Different approaches have been used to automate the web usability assessment (Scholtz et al. 1998; Hong et al. 2001; Ivory, 2001; Paganelli and Paterno, 2002; Paterno, 2003). These were summarised in Chapters 2 and 3 and significant weaknesses identified:- lack of intelligence, and little use of customisation. Several tasks must be achieved as part to the customisation of the usability measure for web pages. These include computing web metrics, reporting the web page usability evaluation to the webmaster, and learning from the webmaster's feedback for future evaluation. Different learning and intelligent algorithms are used to perform the usability evaluation and its customisation. The intelligence implemented in this study is discussed in Chapter 4 and uses the agent technique to control its application.

The first section of this chapter is a description of the web aspects investigated. In the second section there is a discussion of the web metrics calculated and the third section involves the research sample and the HTML tags included, followed by a detailed description of the agent methodology.

6.2 Web Aspects Under Scrutiny

As mentioned in Chapter 2, there are large numbers of web design recommendations, which provide web developers with useful usability guidelines (Nielsen, 2000; Sklar, 2000; Ivory, 2001; Ivory and Hearst, 2001; Lynch and Horton, 2001; Nielsen and Tahir, 2002). This abundance of usability guidelines only applys to design guidelines for English web pages but most might be also appropriate for web pages constructed in similar Western languages. However, these guidelines may not be reliable for languages with different writing directions, different characters and scripts such as Arabic.

One of the main goals of this research is to establish some basic usability guidelines for maintaining legible Arabic web pages while, at the same time, validating the current usability guidelines maintained for English web pages. Therefore, the web aspects investigated by this research are those considered the main factors for web page legibility. This discussion will cover:

- The text and the background colour contrast. Web experts always recommend the use of high colour contrast between the text and the background since the colour combination can either promote or deter the web page readability (Rigden, 1999; Nielsen, 2000; Nielsen and Tahir, 2002).

- The number of colours used in the design of the web page. Since too many colours in the design reduces their functionality which in turn affects the readability (Badre, 2002).

- The length of the browser title because it increases the load on the users' memory. Minimising the use of the users' memory is one of the major usability recommendations (Nielsen, 2000).

- The Number of links in the web page. Many links within the web page might distract users from the main web page content (Nielsen, 2000).

- Changes in the font sizes. The font size has a major influence on the legibility of web page (Laarni, 2003).

- The font effects such as: bold, italic and underlined.

- The alignments of both text and headings contained in the web page. The starting point of the text affects readability; see the current usability recommendations mentioned in Chapter 2 Section 2.7.

- The number of images in the web page and their size. There are several usability recommendations that advise the developer to minimise the number of images in the web pages unless they are necessary.

It is obvious that almost all of the different web aspects mentioned previously were related to the presentation of the text in the web page. The text specification investigated in the research included the font size, colour and position. Other factors that affect the font appearance include bold, italic and underline. Additionally, other aspects, that might reduce the scanability and readability, such as the number of links and the number of images in the web page, are included in the research. However, neither the background image nor the images' contents are included in the current analysis.

6.3 Web Page Metrics Investigated in the Study

All the web metrics were chosen to identify common key features that impact on the legibility of the web page. Fifteen web metrics were selected and computed in this study to measure the web aspects under scrutiny. These web metrics were studied by different authors (Ivory 2001; Ivory and Hearst 2001; Ivory and Hearst 2002; Chevalier, Ivory 2003). All of the web metrics selected in this research were used in WebTango, which is one of the automated web usability evaluation tool. However, WebTango processed these web metrics differently. WebTango is based on a statistical model that compares the web page tested with highly rated web pages and the interpretation of the data is different and so is the usability output rate. Thus, the WebTango output cannot be compared with the usability rate discussed in this thesis.

Table 6.1 summarises the web metrics computed in the study. This research includes several web metrics such as: the number of words in a web page, number of font sizes, number of colours, number of links, etc.; these web metrics are just examples of quantifying web criteria. In other words, these web metrics can be extended to involve different text effects, colour usage, or the font size and the position of a sentence in the web page. All of these considerations should be measured in the future researche.

6.4 Research Sample and Targeted HTML Tags

The aim of this research is to construct an agent that analyses and evaluates all the above web aspects for both Arabic and English web pages. The Java language was chosen to construct the agent because of its ability to execute downloaded code on remote hosts in a secure manner (Arnold et al. 2000). This ability makes it suitable for a network environment like the web. Also, it increase code productivity because of its ease of programming, safety and multithreading features (Arnold et al. 2000).

Web Metric	Description
Web page size	The actual size of the web page in bytes
Word count	The number of words on a web page
Bold words	The number of bold words on a web page
Italic words	The number of italic words on a web page
Underlined words	The number of underlined words on a web page
Number of links	The number of links on a web page
Heading alignments	Heading alignments employed and their number on a web page
Text alignments	Text alignments employed and their number on a web page
Number of images	The number of images on a web page
Images sizes	The total of images sizes within a web page in bytes
Words in title	The number of words in a web page browser title
Font sizes count	The number of font sizes used in a web page
Font size	All the font sizes employed in the web page
Text and background colours	The colours employed in the web page
Colour count	The number of colours used in a web page

Table 6.1: Web Page Metrics Computed in the Study

However, the multilingual interface agent's implementation was based on HTML source code for both Arabic and English web pages. In other words, the agent inspects the web page design based on its HTML language layout. The selection of HTML was based on several factors; it is a very popular language, it is widely used to

construct Arabic and English web pages, is accepted by different browsers and is one of the official web languages endorsed by (W3Consortium 2004a; W3Consortium 2004b). Furthermore, Java has a HTML parser library that can simplify the analysis of the HTML documents.

Several HTML tags were used to identify the targeted web aspects. The following HTML tags are under scrutiny:

- For the colour contrast evaluation, all the tags and those of their attributes that can change either the background or text colours were considered in the analysis of the HTML document. For instance, the Java HTML parser built to analyse changes in the background and text colour relied on the *body*, *table* and *font* HTML tags.
- Only one tag can determine the browser's title.
- The anchor tag was also considered in the HTML parser analysis because it participates in determining the number of the links in the web page.
- The font size attribute can be analysed from both the *basefont* and *font* HTML tags.
- Other font effects such as bold, italic and underlined can be analysed by *B*, *I* and *U* HTML tags respectively.
- Both the headings and text alignments can be analysed from several HTML tags: *body*, *col*, *div*, *h1*, *p*, *table* and *td*.
- All the images in the web page tested can be examined from the *img* HTML tag.

The determination of all these HTML tags is a crucial step for the web metrics computation. The HTML parser approach is discussed in the next section.

6.5 The Automated Web Usability Evaluation Methodology

The automated web usability evaluation approach utilised in this research consisted of four individual but related phases: the web page analysis phase, the web page evaluation phase, the learning phase, and the web usability evaluation controller. This approach was developed to evaluate the Arabic and English web pages that mainly contained information. In other words, the usability evaluation for both languages was based on the web page layout but it did not include any analysis of web page content such as: spelling, grammar, structure of the sentences, etc.

Figure 6.1 illustrates the architecture of the multilingual interface agent and shows its four phases and how they cooperate with each other. The cooperation is shown by the information passed between them. Each phase is dependent on the phase before, which provides it with the data necessary to perform its own tasks, and all of them cooperatively work together to achieve the agent's goal. The agent's major goal was to customise the web usability evaluation to the webmaster's preferences. The webmaster is presented as the owner of the automated usability multilingual interface agent.

The evaluation process is started when the webmaster inserts the URL address, for the web page to be tested. The four different, distinct phases follow the downloading of the HTML source code. A detailed description of each phase is presented in the next sections.

6.5.1 The Web Page Analysis Phase

Downloading the HTML source code activates the analysis phase, which can be considered as the agent gateway for the web page evaluation. Two major tasks are performed by the analysis phase. The first task is to identify the HTML tags concerned in the study within the web page tested. The second task is to compute the relevant web metrics. The HTML parser and the Web Metrics Computation perform these two tasks respectively.

A HTML document is divided into tags, text and comments. But the later are not included in this study since they do not affect the web page layout and the study also does not yet consider the other layout effects such as Java scripts, CSS, etc.. The HTML parser identifies the HTML tags mentioned above, and classifies them into *start, end* and *simple* HTML tags, which simplifies the identification of the important tags. These tags and their corresponding attribute(s), if there are any, can be used to collect some quantitative data from the HTML document. Therefore, the identified tags are considered as delimiters necessary for the web metrics computation section.

The web metric computation applies to the text, identified by the tags. The actual metric is often a word count but for some tags it can be a byte count or repetition count, etc.. For instance, in order to determine the length of the browser title, the HTML parser identifies the existence of the beginning and the end of the title tag and then passes delimiters to the web metrics computation section. The latter section associates the delimiters with the text between them and counts the actual words. For the simple tags, the HTML parser passes only one delimiter to the web metrics computation section, which starts its calculation instantly. All the web metrics calculated are passed on to the web usability evaluation controller phase.

6.5.2 The Web Page Evaluation Phase

The raw data for this phase are the web metrics computed from the previous phase. A fuzzy logic model has been used to control the evaluation decisions and to infer the usability rate from the impact of different web metrics. The usability rate reflects the quality of the web page design. In other words, usability rate indicates the degree of effectiveness, efficiency and users' satisfaction for different web aspects. The higher the usability rate the better the usability weight and utilisation of different web aspects. The descriptions of the web metric measures vary consequently; four fuzzy input sets types are used to measure the following:

1- The length of a specific web aspect. These fuzzy input sets are used for length quantities such as browser title length. This type of fuzzy sets is divided into nine levels and described as: *VeryShort, Short, MildShort,*

Shortish, *Medium*, *Longish*, *MildLong*, *Long* and *VeryLong*.

2- The representation of specific web aspects. For instance, text specification (bold, italic, underline) web metrics which are measured by their number of words, are described using nine fuzzy sets: *VeryFew*, *Few*, *MildFew*, *Fewish*, *Medium*, *Manyish*, *MildMany*, *Many*, and *VeryMany*. The same fuzzy set can be applied for web metrics such as the number of links in the web page and image sizes.

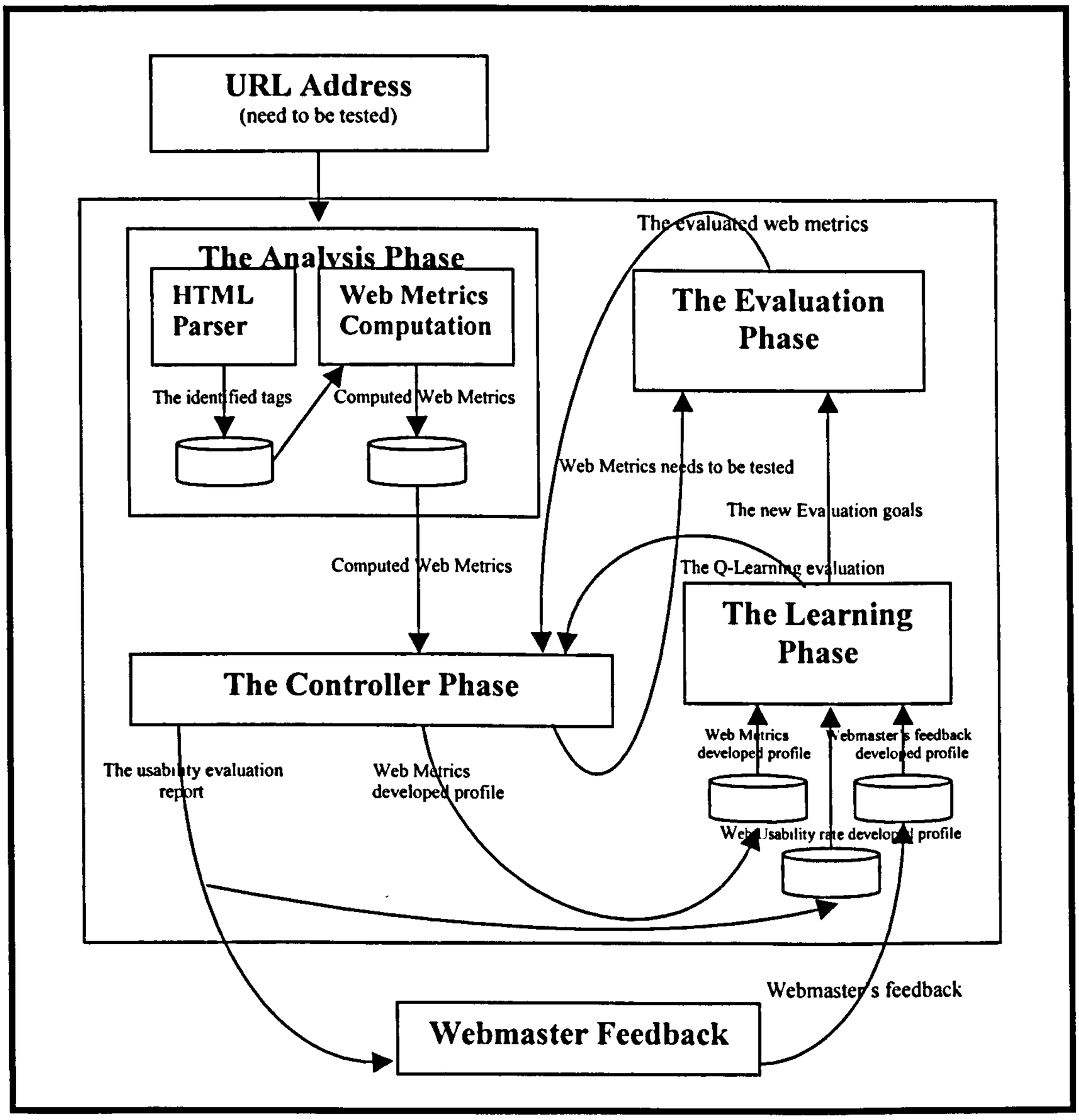


Figure 6.1: The Multilingual Interface Agent Architecture

3- The size of different web aspects. For example, the font size fuzzy input set are also combined from nine fuzzy inputs *Verysmall, Small, MildSmall, Smallish, Medium, Largeish, MildLarge, Large, VeryLarge*.

4- The contrast of different web aspects. The most important contrast in the web page should be between the text and the background colour. The fuzzy input sets for measuring the colour intensity, for both the background and the text colour, are described as: *Lightest, VeryLight, Light, MildLight, Lightish, Medium, Darkish, MildDark, Dark, VeryDark* and *Darkest* (Abulkhair and North, 2003).

Despite the four types of the fuzzy input set used, all of them produce the same fuzzy output presented in the following set: *VeryLow, Low, MildLow, Lowish, Medium, Highish, MildHigh, High* and *VeryHigh*.

Different fuzzy rules have been applied in order to tune the fuzzy input to produce a proper fuzzy output, representing the usability rate. These fuzzy rules are derived from previous studies together with the results from the user satisfaction conducted as part of the research. For instance, to determine the usability rate for one of the bold text specifications, the fuzzy rule is:

IF (BoldText IS VeryFew) THEN (UsabilityRate IS VeryHigh)

Many web experts recommend using some bold text (Nielsen, 2000; Nielsen and Tahir, 2002), and at the same time they recommend avoiding underlined text (Nielsen, 2000; Lynch and Horton, 2001). The user satisfaction validated both these recommendations. Therefore, the corresponding underlined text fuzzy rule differs and it is of the form:

IF (UnderlinedText IS VeryFew) THEN (UsabilityRate IS MildHigh)

However, the font size fuzzy rules differ from these rules. One form of the font size fuzzy rules is:

IF (FontSize IS VerySmall) THEN (UsabilityRate IS Low)

The fuzzy rules implemented to measure the length of specific web aspect, such as the length of the web browser's title within the web page, is:

IF (TitleLength IS VeryShort) THEN (UsabilityRate IS VeryHigh)

All the fuzzy rules mentioned here are samples to illustrate the fuzzy evaluation model, but the model is not limited to them. A large number of fuzzy rules for different fuzzy input set types which vary according to the web metrics tested are needed. All the fuzzy rules were derived from the current usability guidelines (see Chapter 2) in combination with the user satisfaction explained in Chapter 7 and empirical experiments conducted by the researcher. After experimenting with different types of membership functions to calculate fuzzy input, it was clear that the triangular function gave the most consistent answers; so a triangular function was used to calculate the fuzzy membership for all the fuzzy input types. The crisp value,

produced from the defuzzification process, varies from (0 – 100) where 0 is the lowest and 100 is the highest usability rate.

Colour contrast was one of the most complicated web aspects measured, and its fuzzy model is composed of two nested fuzzy phases (see Appendix *C*). The first fuzzy phase was to measure the intensity of both the text and the background colours, whereas the second fuzzy phase was to measure the colour contrast between the text and background intensities. The details of both phases are presented in the next sections.

6.5.2.1 Colour Intensity Phase

It is necessary to determine the colour intensity of both the text and the background colours prior to the colour contrast assessment, which in turn is necessary to measure the usability rate. This phase uses the fuzzy rules to infer the colour intensity from its RGB components and their variations. Each of these components can be split into five fuzzy sets: *Dark*, *MidDark*, *Medium*, *MidLight* and *Light*, according to the amount used by the web designer and then combined to give an overall colour intensity measure. To determine the overall colour intensity various fuzzy rules are applied to combine the RGB components in order to give a specific colour intensity value. The colour intensity is quantified as: *Darkest*, *VeryDark*, *Dark*, *MildDark*, *Darkish*, *Medium*, *Lightish*, *MildLight*, *Light*, *VeryLight* and *Lightest*. The ranges of each element in the colour intensity fuzzy set are presented in Figure 6.2. The fuzzy rules involve combining the RGB intensities determined earlier in the fuzzy sets to produce a fuzzy colour intensity. This process can be illustrated by the following example:

IF (Red IS Dark) AND (Green IS Dark) AND (Blue IS Dark) THEN (ColourIntensity IS Darkest) (1)

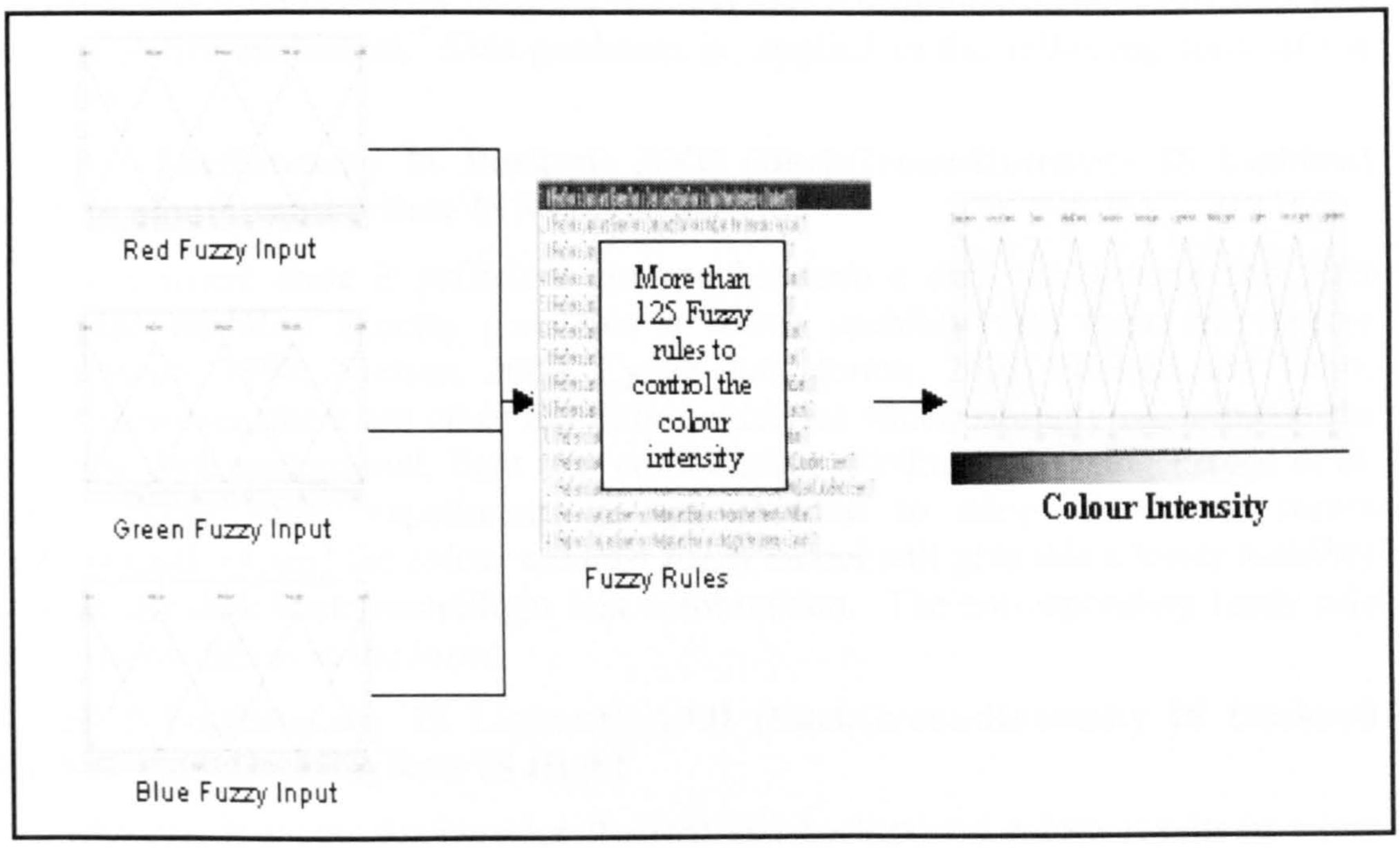


Figure 6.2: The Intensity of the Colour

Having established a general rule for colour intensity measuring, we also considered the special case of red/green colour blindness which is fairly common. The red/green colour blindness case is dealt with using the RGB colour intensities to produce an intensity of either Red or Green. This fuzzy rule is of the form:

IF (Red IS Dark) AND (Green IS Medium) AND (Blue IS Dark) THEN (ColourIntensity IS Dark) AND (ColourBlind IS Green) (2)

There are around 125 fuzzy rules to derive the colour intensity measure, and all of them followed the form of either example, (1) or (2). Both forms of fuzzy rule are based on experimental results.

The colour intensity crisp value, computed from the defuzzification process, varies from (0-100) where 0 is the darkest and 100 is the lightest.

6.5.2.2 Colour Usability Rate Phase

Once the colour intensities have been established they are employed in the colour usability rate phase. This is a crucial part of the model, because it integrates the two colours intensities with the most common colour blindness case to produce an appropriate colour usability rate. Thus the specification of the colour intensity fuzzy set is inherited from the previous phase to infer the colour usability rate from the colour contrast measure. The colour usability rate can be categorised as *VeryHigh, High, Highish, Medium, Lowish, Low* and *VeryLow*. The *VeryHigh* colour usability rate is achieved by the most widely recommended usability guidelines for colour. For instance, the combination of a white background together with black text (but not the other way around) will be given the highest colour usability rate, (Shneiderman, 1998; Nielsen, 2000; Lynch and Horton, 2001; Nielsen and Tahir, 2002) whereas a low contrast between the text and background colour indicates low colour usability evaluation. This guideline is applied in the following form of the fuzzy rules:

IF (TextIntensity IS Darkest) AND (BackGroundIntensity IS Lightest) THEN (ColourUsabilityRate IS VeryHigh)

Even where there is sufficient contrast between a dark background and light text, some usability experts give this a lower usability rate than the reverse (Shneiderman, 1998; Nielsen, 2000; Lynch and Horton, 2001; Nielsen and Tahir, 2002). However, there are other usability guidelines which strongly recommend the use of the dark background, light text combination (Rivlin et al. 1990; Preece et al. 1994). After some experiment, it was decided to adopt the more recent recommendations, and the colour contrast fuzzy model will give this a lower usability rate than the dark background/light text combination. The corresponding fuzzy rule used to assess this is in the form:

IF (TextIntensity IS Lightest) AND (BackGroundIntensity IS Darkest) THEN (ColourUsabilityRate IS High)

whereas, low contrast between the text and background colour results in a low colour usability evaluation. This gives us a group of rules as follows:

IF (TextIntensity IS Dark) AND (BackGroundIntensity IS Dark) THEN (ColourUsabilityRate IS VeryLow)

IF (TextIntensity IS Light) AND (BackGroundIntensity IS Lightest) THEN (ColourUsabilityRate IS VeryLow)

IF (TextIntensity IS Medium) AND (BackGroundIntensity IS Darkish) THEN (ColourUsabilityRate IS VeryLow)

There are more than 140 fuzzy rules in the model to cover different aspects of colour contrast. These fuzzy rules are combined to give a reliable colour usability rate, and all the fuzzy rules are illustrated in Figure 6.3. In this figure different intensity values from the text and background colours can give rise to different usability rates. In other words, the right hand section of the surface in Figure 6.3 represents the light text over dark background colour combination, which is rated lower than the opposite colour combination presented on the left surface section. The crisp value, produced from the defuzzification process, again varies from (0 – 100) where 0 is the lowest and 100 is the highest colour usability rate.

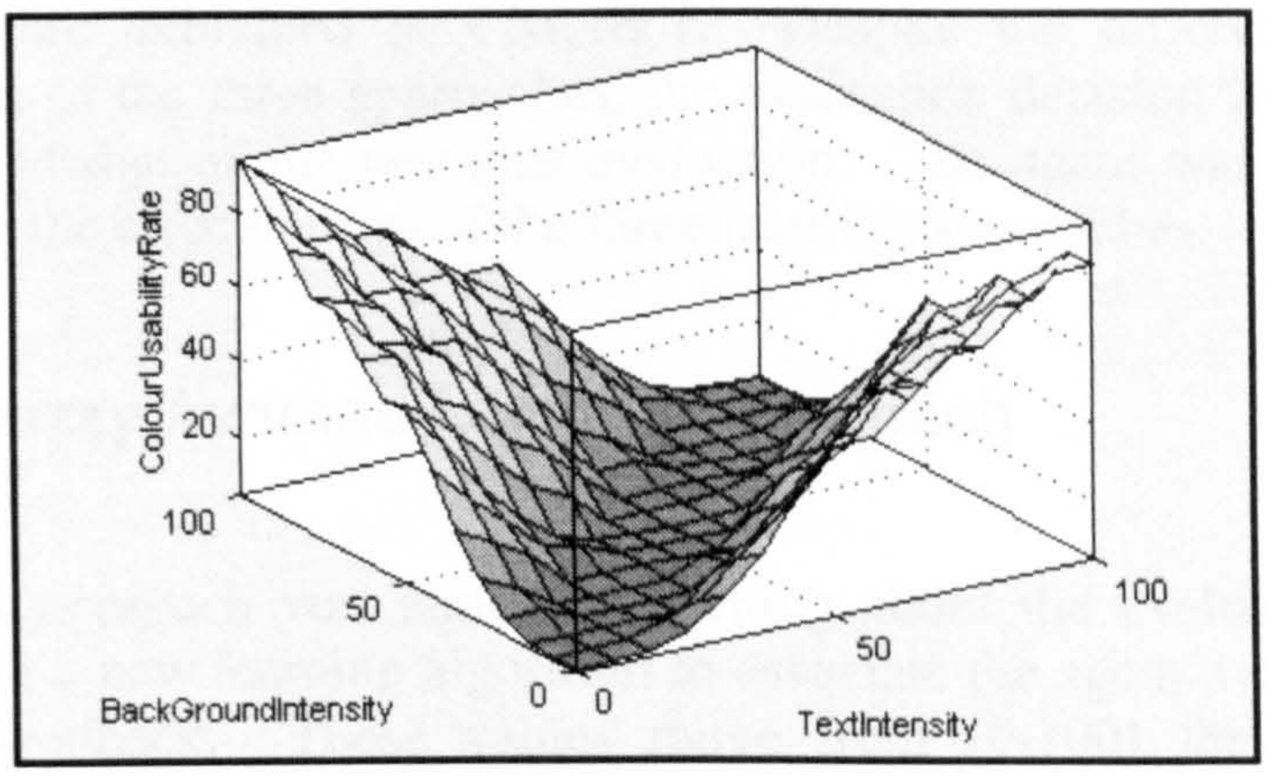


Figure 6.3: Colour Usability Rate [the dark colour shaded point to the lower colour usability rate, where the two upper edges to the right and left indicate high usability rate]

Figure 6.4 illustrates the algorithm used to construct all the fuzzy models. It is obvious that all of them started to fuzzy the input, which represented a specific web metrics need to be evaluated. Then the algorithm ended by defuzzifying the output and converting it to a crisp value.

- Fuzzify the web metric input
- Infer the fuzzy output by processing the fuzzy input into fuzzy rules
- Defuzzify the output

Figure 6.4: Fuzzy Model Algorithm

All the fuzzy models evaluation results are passed to the controller phase as shown in Figure 6.1. This figure also shows that the Q-learning web metrics evaluated are sent from the learning phase to the controller phase.

6.5.3 The Learning Phase

In order to customise the agent's web usability evaluation, it is assumed that the agent would get a feedback from the webmaster each time it is given a URL address to test. Each web metric tested should have either positive or negative feedback, which are important in determining and customising the next evaluation. Three learning approaches were used to provide this aspect of the agent's intelligence: the fuzzy average based approach, the fuzzy learning approach and the Q-learning based approach. Several reasons stood behind selecting more than one learning approach, these reasons are explained in Chapter 1. Despite the differences between the implementation of the three approaches, the evaluation decision for all of them was driven by knowledge of the previous evaluation. The agent was run about twenty times to realise the effectiveness of the three learning approaches.

6.5.3.1 The Fuzzy Average Learning Approach

The first approach was aimed at learning about the evaluation of each web aspect based on a new learning algorithm to associate the agent's evaluation with the webmaster's feedback. These values range from (0-100) for each web aspect evaluation. For each web aspect the agent's learning algorithm calculates the average of both the webmaster's feedback and the web aspect evaluation. Both the usability evaluation average and the webmaster's feedback average are changed to fuzzy inputs, and then they integrated in a fuzzy model to update the fuzzy rules mentioned in the evaluation phase. For example, when testing the browser's title length, the first time, the agent evaluates the title by using the fuzzy model described above. Then each subsequent time the learning algorithm of the agent calculates the average of the web browser's title evaluation and the webmaster's feedback. The new evaluation varies as result of the webmaster's feedback; it might increase or reduce the usability rate.

Figure 6.5 illustrates the average learning algorithm steps, which are run for each web metrics tested. As can be seen from Figure 6.5, in the third and fourth steps both averages are converted to fuzzy inputs, which are integrated in an updated fuzzy model. The webmaster's feedback is described using five fuzzy input sets: *StronglyAgree*, *Agree*, *Neutral*, *Disagree* and *StronglyDisagree*. Whereas, the usability evaluation is represented by nine input sets: *VeryHigh*, *High*, *MildHigh*, *Highish*, *Medium*, *Lowish*, *MildLow*, *Low* and *VeryLow*. Forty five fuzzy rules were implemented to infer a proper usability evaluation update. The customisation was introduced simply by converting the evaluation into *VeryHigh* or *High* if the webmaster's feedback is *StronglyAgree*. To have consistent evaluation conversion, each webmaster's

feedback is assigned two ranges of the usability rate fuzzy sets. For instance, for the *Disagree* feedback there are two fuzzy usability rate sets, *Lowish* and *MildLow*. The determination of the alternative usability rate is based on the average of the previous usability rate. In other words, the *StronglyAgree* and *Agree* webmaster's feedback always pointed to positive feedback and then high usability rate; while *Disagree* and *StronglyDisagree* indicate the negative feedback; in that case they will drag the usability rate down to the minimum. *Neutral* feedback stabilise the usability rate to medium.

The following are some of the fuzzy rules, which can be considered as examples of all the fuzzy rules and summarises the evaluation conversion mechanism:

IF (UsabilityRateAvg IS MildHigh) AND (WebmasterFeedbackAvg IS StronglyAgree) THEN NewUsabilityRate IS VeryHigh

IF (UsabilityRateAvg IS MildHigh) AND (WebmasterFeedbackAvg IS Agree) THEN NewUsabilityRate IS MildHigh

IF (UsabilityRateAvg IS MildHigh) AND (WebmasterFeedbackAvg IS Neutral) THEN NewUsabilityRate IS Medium

IF (UsabilityRateAvg IS MildHigh) AND (WebmasterFeedbackAvg IS DisAgree) THEN NewUsabilityRate IS Lowhish

IF (UsabilityRateAvg IS MildHigh) AND (WebmasterFeedbackAvg IS StronglyDisAgree) THEN NewUsabilityRate IS Low

The fuzzy average learning rules infer a new usability rate for each web aspect tested. The crisp value of the new usability rate ranges from (0-100), where 0 is the lowest and 100 is the highest. The next usability evaluation would be affected by the results from this phase, which would be passed to the evaluation phase.

Repeat the following for each web metric separately

- Evaluate the web metric using the fuzzy model
- Receive the webmaster feedback
- Calculate the webmaster's feedback average and convert it to fuzzy input
- Calculate the web metrics' evaluation average and convert it to fuzzy input
- Update the evaluation:
  new web metric's evaluation = updated evaluation

Figure 6.5: The Fuzzy Average Learning Algorithm

6.5.3.2 The Fuzzy Learning Approach

This learning approach is quite similar to the previous learning approach in using a fuzzy logic. However, this fuzzy approach is based on the webmaster's feedback instead of the webmaster's feedback average and the previous usability rate average. The new usability rate is determined on a basis of webmaster's feedback from the previous usability evaluation. The fuzzy input sets are *StronglyAgree, Agree, Neutral, Disagree* and *StronglyDisagree*, which ranges from (0-100) and represent the webmaster's feedback. The usability rate fuzzy output is divided into nine sets, as in the previous method, which are *VeryHigh, High, MildHigh, Highish, Medium, Lowish, MildLow, Low* and *VeryLow*. To produce a reasonable usability rate five fuzzy rules are involved in the determination, for example:

IF (WebmasterFeedback IS StronglyDisAgree) THEN NewUsabilityRate IS ranges between (Low and VeryLow)

Webmaster's Feedback	The New Usability Rate Range
StronglyAgree	*VeryHigh, High*
Agree	*MildHigh, Highish*
Neutral	*Medium*
DisAgree	*Lowish, MildLow*
StronglyDisagree	*Low, VeryLow*

Table 6.2: The Ranges for the New Usability Rate Produced by the Fuzzy Learning Approach

Table 6.2 summarises how the webmaster's feedback can affect the determination of the new usability rate. The usability rate crisp value ranges between (0-100), where 100 is the highest usability rate. The new usability rate calculated by this approach can affect the next evaluation.

6.5.3.3 The Q-Learning Approach

The third learning approach was based on the Q-learning approach which is a standard of reinforcement learning method (see Chapter 4 Section 4.9.1). The Q-learning approach is founded on three major components: the quantity of the web aspect tested (state), the state of the usability evaluation (action), the webmaster's feedback (reward). Each time the new Q is calculated by the following Q-function:

$$Q(s, a) = Q(s, a) + \alpha*(r + \gamma*maxQ - Q(s, a))$$

Where s is the state of the web aspect tested, and it represents the actual quantity of the web aspect needs to be evaluated. For example, when testing the number of links in a web page the state would vary from *verymany* to *veryfew*. In other words, each web aspect can be in one of the nine different states determined as fuzzy input. The a is the action of the current state, and it represents the appropriate usability evaluation for the state tested. Each action can be chosen based on the *Softmax* action selection technique, which gives actions with higher estimated values higher probabilities of being selected. The r is the reward value and it represents the webmaster's feedback for the action taken. The α is the learning rate of the agent, and it is set between (0 - 1). The γ is the discount factor to present the value of future rewards; it is also set between (0 – 1). The maximum Q value can be determined by choosing the action with the highest webmaster's feedback reasonable for the state tested.

All this refers to the current state of evaluation and totally neglects the next state estimation. Each state was considered as a terminated state. All the Q-values were initialised by zero and changed as the state was tested. The value of alpha α is equal to 1, which is the highest learning rate. While gamma γ value is equal to 0.1 to maximise the immediate rewards. Figure 6.6 shows the Q-learning algorithm steps.

The *Softmax* technique is used to rank and weight all the values according to their values estimates, which are identified on a basis of varying the action probabilities (Sutton and Barto, 1998). The Softmax action selection technique is calculated using the following probability formula:

$$\frac{e^{Q(a)/\tau}}{\sum_a e^{Q(a)/\tau}}$$

Where a describes the actions, and the τ is a positive parameter called the *temperature*, which is selected to be equal to one. A high *temperature* value causes all the actions to be equi-probable, unlike its value when it becomes nearly zero. A greater difference appears in the selection probability for actions that differ in their value estimates (Sutton and Barto, 1998).

The algorithm steps are as follow:

- Initialise the Q- values, Q(s, a)
- Observe the current state, s
- Choose an action by using Softmax action selection technique
- Determine the reward, r received by the agent in the s state
- Calculate the new Q-value
- Update the Q(s, a) with the new Q-value

Figure 6.6: The Q-learning Algorithm

Using the state value identified from the web page tested and the action selected, all the Q-values are passed to the controller phase.

6.5.4 The Web Usability Evaluation Controller

The controller section was constructed to manage the movement of the data between different phases of the agent. It receives the web metrics computed and prepares for the evaluation from the analysis phase. Then it distributes them to two different phases of the agent, the web usability evaluation phase and the learning phase. Both phases return their evaluation to the controller, which passes the two phases evaluation report to both the webmaster and the learning phase.

Three different profiles were developed in this phase, the web metrics, webmaster's feedback and the usability rate profiles. The usability rate profile consisted of the usability rate for the last web metrics evaluated in addition to the usability rate for all the previous assessments. Also the webmaster's profile contains the last feedback in addition to the collection of all webmaster's feedbacks received. These various profiles can be used by different learning approaches. Figure 6.1 illustrates the input and output flow of the data from, and to, different phases.

6.5.5 Webmaster's Feedback

After the evaluation has been completed, the controller phase of the agent passes the evaluation report to the webmaster. Then the webmaster replies to the agent with feedback in the form of numbers which reflect the webmaster's acceptance or rejection of the usability rate produced by the agent. The webmaster's feedback ranges from (0 to 100) where 0 was the lowest feedback and the 100 was the highest. As soon as the agent gets the webmaster's feedback, it develops the webmaster's profile. This profile is considered a crucial step prior to the customisation performed by the learning phase of the agent.

The webmaster's feedback can be simulated by giving random numbers to the agent instead of relying on a real webmaster. In this way, the agent gets unbiased, lowers cost feedback more quickly, which affects its productivity. Furthermore, agent's customisation can be repeatedly and consistently tested against various webmaster's responses, ranging from similar to drastically different feedback.

The webmaster's feedback simulation is divided into four ranges each with five iterations. In the first set the webmaster's feedback ranges between *StronglyDisagree* and *Disagree*. Whereas in the second set the webmaster's feedback vary between *Disagree* and *Agree*. The third set of the webmaster's feedback ranges between *Agree* and *StronglyAgree*. The last set is reflected the stable condition in the feedback, when it is in the same range as the third set of the webmaster's responses.

6.6 Conclusion

This chapter presented the multilingual interface agent architecture and how it can be implemented using agent methodology. The discussion involved the evaluation process and the implementation of the agent's intelligence, which involves three learning approaches. Unlike other automated web assessment techniques, the multilingual interface agent approach customises the assessment to reflect the webmaster's preferences. The idea of customisation of the evaluation supports the dynamic web usability guidelines.

The results from the evaluation phase will be discussed in Chapter 8, whereas the details of the three learning approaches and the comparison between them will be discussed in Chapter 9. The next chapter presents the user satisfaction findings and identification of the basic usability guidelines appropriate to evaluate the Arabic and English web pages.

Chapter Seven

Linking Web Usability Guidelines with the User Satisfaction

7.1 Introduction

The WWW is now the largest information collection and communication tool, and it attracts people from all over the world. The WWW is no longer restricted to displaying and interacting in English as it used to be because the number of the non-English web users is expected to increase to 70% by 2005 as (IDC, 2001; Sun, 2001) indicate. In order to break the barrier between different WWW services and the non-English speaking users, web service localisation has been used to customise the services for different languages. The localisation process is performed on two levels; the surface level that involves data translation, amended punctuation and other conventions familiar to the targeted users and the cultural level (Sun, 2001). The cultural level means the appearance adjustments which include the selection of images, colours and communication patterns to conform the targeted audience's tastes (Russo and Boor, 1993; Sun, 2001). Therefore, web developers should consider many localisation aspects in addition to data translation, because different languages mean different cultures and values. It is this aspect of customised web evaluation which is discussed in this chapter.

One of the weaknesses of the existing usability guidelines is that they have been designed for, and evaluated on, Latin scripts. No guidelines appear to exist for other scripts. The first difficulty in assessing non-English web pages is the lack of any reliable guidelines. In order to explore multilingual evaluation the researcher first had established guidelines and chosen Arabic to explore. By performing user satisfaction for Arabic and English in this way, the usability guidelines should more reliably meet the users' needs.

This chapter discusses the usability testing conducted to localise the usability guidelines and to determine the similarities and the differences between different cultures, in particular Arabic and English. The usability testing can be considered as a filtering mechanism for the current usability recommendations to validate these guidelines as being reasonable for local users.

7.2 Usability testing and web design

In many studies usability testing is used to provide an insight into different problems that might prevent the users from fulfilling their goals (Parush, 2001). It is considered the closest technique to the users, i.e., through the interaction with users, the tester can understand the users' needs and preferences. It is hard to determine these aspects by using other usability evaluation techniques (Law and Hvannberg,

2002; Rohn et al. 2002). Therefore, user testing is widely used in studying the impact of web design on users and recognising cultural effects.

Localizing web pages is not limited to translating the content of the web pages as mentioned earlier; instead, it should include other considerations, such as: layout, interaction, symbols and colours used in the web pages (Russo and Boor, 1993). Sun conducted a study that used the user testing to learn how cultural markers affect web usability. That study was based on four main cultural markers: language, colour, images and page layout (Sun, 2001). He found that participants unconsciously applied their cultural preferences to evaluate the design of the web pages (Sun, 2001).

User testing is performed over a text design and the results show that significant insights can be gained as a result (Gee, 2001). This study is aimed at determining the foundations for comfortable, easy reading, and its user testing is divided into three distinct issues: interface issues, reading behaviour issues, and reader attitude issues (Gee, 2001). Goonetilleke, Duffy and Jacques (2001) implemented a different form of user testing to understand the higher-level factors of usable web page design and determine which of them is the most important to users. The participants evaluated the web sites based on the sixteen different issues relevant to readability.

Furthermore, user testing has been used to measure the interaction with different aspects in the web page design from the cultural view. Lee (1999) utilised the user testing to measure cultural effects on user interface design, distributed over fifteen countries. This study involves different objectives relevant to the users and their cultures which encompass understanding the relationship between cultural characteristics and user interaction in addition to identifying cross cultural differences in interaction.

Many interface problems have been established by using usability testing. In fact there are several factors that can affect the users' viewpoint and evaluation of web pages, as discussed earlier, and these factors should be considered in establishing the usability guidelines. Cultural effects are widely recognised as one of the significant factors that affect the web design, because users' preferences differ from one culture to another. Additionally, users' viewpoint and assessment of the web design are influenced by users' knowledge and values, which are also affected by culture.

Since user testing shows significant differences in studies of the cultural effects on the web design (Lee, 1999; Sun, 2001; Shin et al. 2004), it is an effective way to empirically validate the existing usability guidelines over the Arabic and English cultures. In the next sections there follows a detailed discussion about how culture might affect usability.

7.3 User Satisfaction Design

User satisfaction was applied to examine users' preferences in relation to the current usability guidelines; the details of the user satisfaction were mentioned in Chapter 5. A total of 196 participants completed the three Arabic questionnaires whereas the English questionnaires were completed by 94 participants. They went through several questions related to various web design specifications; additionally they rated some web design examples according to their preferences. English and

Arabic participants evaluated 34 web pages, each focused on different usability features. The user satisfaction was developed using HTML forms with an easy click buttons and menu choice selection to attract the users and simplify the form filling process.

7.4 Testing Examples and Tasks

The current study used six questionnaires; three in Arabic and three in English, each of them was associated with different usability issues. However, the structure of the questionnaires was the same, they all began with a summary of the main issue of the questionnaire, followed by several multiple-choice questions. Each questionnaire ended with message of appreciation.

The multiple-choice questions were presented to build up knowledge about how to measure the web design, and what features the web pages should include to be more readable and attractive. These questions were not limited to the specification of the web design but they were extended to include users' preferences for some aspects. For instance, the participants were asked about colour usage and whether they liked subtitles and important text included in the web pages to be in different colours. On the other hand, some other questions investigated users' scanning habits and the subtitles they preferred. The meaning of different colours and combinations of colours was also explored to build knowledge about the impression the colours can add. Some of the questions involved online-examples which were rated by the participants.

Each example in the questionnaires examined certain usability issues about users' preferences and views on legibility for both English and Arabic users. Each participant had a chance to rate the content of these examples, on a scale of three levels. Testing examples were developed using HTML and were linked to the questionnaires' body by using an iframe HTML tag. This made it easier for the users to remain in the questionnaire web page without going back and forth when evaluating the examples. Figure 5.2 shows one of the examples used in the grading questions. The examples were displayed in the same order in both languages.

7.4.1 Participants

The main feature for remote usability testing is that the test can be distributed to any place in the world. Direct contact with users is not necessary for the participants to complete their jobs. No restrictions were applied to participate in questionnaires, i.e. the questionnaires' replies were not limited to certain group of people or gender; however, this study did not include children. Participants were not required to learn any special techniques, nor did they have to initialise or download a specific internet tool. Normal internet users with ordinary skills to work with the WWW were welcome to join the participants, but they needed to be either Arabic or English fluent speakers in order to answer the question.

All the questionnaires were distributed via email using different email lists for British, Arab students and staff within the University of Sheffield. Additionally,

different email lists were used to distribute the Arabic questionnaires, including University students and other groups. In general the vast majority of the participants were postgraduate and undergraduate students of the University of Sheffield and King Abdulaziz University. The following table shows the different number of replies over the three individual questionnaires.

Questionnaire Name	Arabic	English
Colour Questionnaire	83	34
Text Questionnaire	58	32
Title Questionnaire	55	28
Total Number of Participants	196	94

Table 7.1: Participants and the number of replies

For a particular questionnaire, all the participants answered the same questions in either Arabic or English depending on their native language. The details of each questionnaire contents were discussed in an earlier chapter but the participants had to choose a preferred answer or to rate the enclosed examples according to their attractiveness, fast scanning or comfortable reading.

7.4.2 Data Collection

The final datasets consisted of 290 different cases where each case included participants' responses to the questionnaires. These responses were received via email, and were classified into three different classes: colour, text and title responses. The responses did not involve any participants' demographics because they were not considered in the study, except the users' mother tongue identity. However, these responses were aimed to provide some subjective measures necessary to determine some of the usability guidelines.

The subjective measures included the participants' responses to the multiple choice and rating questions. These involved some measures such as attractiveness, scanability, legibility, readability and the utilisation of both colour and font. The multiple choices participated in characterising some of the web aspects used to construct a legible web page. The web aspects characterisations were based on the users' viewpoints. Some of these characteristics related to the users' scanning habits, the colour utilisation, colour combination preferences, the presentation of the important text and headings, link occurrences and title lengths, whereas, the rating questions contributed to the measurement of the impact of some usability guidelines, presented in web design to the users. These rating questions involved rating the colour attractiveness, colour readability of different colour examples. Additionally, they included the rating of different font specifications such as: bold, italic, underlines, different letter sizes, capitalisation (only in English), heading alignment and the text body alignment (for Arabic).

The rating questions measured different usability aspects as a way to construct a legible and attractive web page. These measures were based on three weight scales, which were mentioned earlier in this chapter, to give the participants the opportunity to choose their preference level. The collected data was monitored to calculate its

frequency, in order to determine the highest priority from the lowest, and then validated these preferences with the usability guidelines. Also, a chi-square statistical test was performed to point out the significant differences between the English and Arabic users' preferences. The details of the data analysis and its relation to the current usability recommendations will be discussed in the next section.

7.4.3 Data analysis

The analysis of the data included the analysis of the subjective measures (attractiveness, scanability, legibility, readability and utilisation of both colour and font aspects), mentioned in the last section. The analysis consisted of the findings of the research questionnaire which investigated Arabic and English users' preferences in web design. These findings were associated with each other to identify significant differences between English and Arabic viewpoints. Furthermore, the analysis involved mapping between the current usability guidelines and the collected data (Arabic and English). This mapping studied the relationship between the users' preferences and the current guidelines to produce validated, empirical and accurate usability guidelines.

The structure of the next section starts with the findings of the colour questionnaire for both Arabic and English. It is followed by the comparisons between English and Arabic findings, and then the mapping of these findings onto the current usability guidelines. The same process is repeated for text and title questionnaire too.

Colours influence web information presentation, because they are not limited to adding aesthetic features to the web design. They can be employed to emphasize specific items and to send signals to users (Abulkhair and North, 2003). Colours combinations preferences might vary from one user to another even in the same culture.

7.4.3.1 Colour Questionnaire Data Analysis for Arabic users

In general, the data analysis from Arabic speakers showed a coherent finding that 75% of Arabic users preferred light background colours to the dark and in particular 68% of Arabic users liked a white background colour with dark-blue text more than any other. Also, the findings collected from rating different examples, strongly confirmed that Arabic users preferred light colours for background and Table 7.2 shows the different ratings given to the selected examples. The colours in the table are arranged according to intensity with blue being the lightest and black the darkest. It can be seen from the Table 7.2 that 55% of Arabic users considered a black background as unattractive, and that was the strongest preference among all the other choices.

The results from asking the users to arrange the same four background colours, mentioned earlier, in order of preferences found that a blue background was the most commonly preferred among Arab users at 80%. The second preference was a pink background 66% followed by green 52% and the least favourite background was the black by 57%. These results were consistent with findings in Table 7.2.

Background Colour	Good	Neutral	Bad
Blue	92%	8%	0%
Pink	55%	42%	2%
Green	8%	47%	45%
Black	23%	22%	55%

Table 7.2: Arabic Users Ratings for Colour Attractiveness

However, the readability measure findings differed from the users' preferences. The least readable colour combination was identified as the green background with 53%, while the example with a black background was given the second worst readability measure with only 39%. The Arabic users' readability measure for the pink and blue backgrounds were both considered good with 57% and 84% respectively.

The colour scanability measure was notable, because the findings indicated that 95% of Arabic users preferred the important text and headings within the web page to be presented in different colours. However, the colour usage showed incoherence in colour aspects' utilisation and Table 7.3 illustrates these values.

Colour Usage	Percentage
Document Title	37
Important text	40
Tables	5
Subtitles	17

Table 7.3: Arabic Colour Utilisation

Some of the colour meanings suggested in the user satisfaction seemed to be relatively close to the Arabic participants' opinions and Table 7.4 shows different colour meanings. It is clear that some of the colours' meanings were significantly affected by Arab users' culture; for example, their support for the purity for white, attention for red, calm for blue and spring for green. The green meaning was associated with paradise, and it considered to be the colour worn in heaven as mentioned in "Holy Quraan" (chapter "Alinsan", aya 21).

Arab users did not consider specific colour combinations as part of their culture, except the colour combinations for some football teams, which is the most popular sport in the Arab world. The findings, showed in Table 7.5, indicated that the smallest percentage was given to the meanings presented.

Colour	Meaning	Percentage
White	Purity	93
	Death	0
Black	Mourning	60
	Disgrace	5
Red	Attention	83
	Joy	10
Blue	Sad	5
	Calm	84
Yellow	Happiness	21
	Bright	49
Green	Envy	2
	Spring	84

Table 7.4: Colours' Meanings Responses of Arabic Participants

Colour Combination	Meaning	%
Red text and white background	Joy	28
	Sport team symbol	23
Blue text and white background	Nationalism	15
	Sport team symbol	33
Green text and white background	Peace	59
	Sport team symbol	13
Yellow text and black background	Bugs	25
	Sport team symbol	22
Red text and green background	Christmas	10
	Sport team symbol	11

Table 7.5: The Reflections of Colours' Combinations Meanings on Arabic Users

Despite the variety of findings in measuring the effect of colours on Arabic users, they could still be utilised to be as a basis for establishing colours usability guidelines suitable for Arabic. The discussion of the colour findings for English users is in the next section.

7.4.3.2 Colour Questionnaire Data Analysis for English users

Colours might have different impact depending on culture. The findings showed that English users remarkably preferred light background colours to dark backgrounds with 65%. Also, 47% of the English users liked a dark-blue text over white background, while 32% of them preferred a black text over yellow background. The rest of the English users, 21%, chose other colour combinations with a dark background colour.

Background Colour	Good	Neutral	Bad
Blue	71%	27%	3%
Pink	38%	44%	18%
Green	24%	41%	35%
Black	9%	6%	85%

Table 7.6: English Users Ratings for Colour Attractiveness

The findings collected from rating different examples proved that English users preferred a light to a dark background and Table 7.6 shows the different ratings scales given to each colour. It is clear from Table 7.6 that highest percentage, on the Good scale, is 71% which was given to a blue background, and it was the lightest of all. The most negative result, 85%, was given to a black background, which was the darkest colour, but both pink and green backgrounds were rated as neutral at 44% and 41% respectively.

The English users' preferences, where they arranged the same four on-line examples with the same background colours, were a blue background with 65% as a first choice. Then followed pink with 50% and a green background was third choice with 59%, while a black background was the least commonly preferred by 68%. The findings from this section were consistent with the results presented in Table 7.6.

However, the readability measure findings split into two groups, either good or bad, with blue and pink backgrounds considered as good by 91% and 85% respectively. On the other hand, black and green backgrounds were rated bad by 44% for the black and 53% for the green. It is clear that the English users are attracted to a light background.

Regarding the scanability measures, 82% of the English users preferred a different colour for the presentation of the important text within the web page. They needed this feature for faster scanning, which seemed coherent with the colour utilisation as part of measuring the web aspects' utilisation. Table 7.7 shows different colour utilisation percentage and it is clear that the highest percentage was given to the representation of the important text.

Colour Usage	Percentage
Document Title	24
Important text	53
Tables	6
Subtitles	9

Table 7.7: English Colour Utilisation

Table 7.8 associates different colours meanings with their findings and it is clear that many of the colours' meanings were relatively well matched to the English users' views. It seemed that some of the colour meanings were strongly affected by English users' culture such as: calm for blue, purity for white, attention for red and bright for yellow. Some of the suggested meanings were considered very popular in Western culture by (Terpstra and Sarathy, 1994), for example, sad for blue and envy for green. But English users gave least support to these meanings and rather selected other meanings suggested.

Colour	Meaning	Percentage
White	Purity	77
	Death	3
Black	Mourning	53
	Disgrace	0
Red	Attention	77
	Joy	9
Blue	Sad	6
	Calm	86
Yellow	Happiness	15
	Bright	77
Green	Envy	18
	Spring	50

Table 7.8: Colours' Meanings Responses of English

Some colour combinations are related to specific occasions such as red and green associated with Christmas in Western culture and so, in English culture. The findings showed that 68% of the English users supported the Christmas meaning for the combination of red text and green background. The other meanings of the colour combinations findings, mentioned in Table 7.9, were not strongly supported.

Colour Combination	Meaning	%
Red text and white background	Joy	9
	Sport team symbol	29
Blue text and white background	Nationalism	15
	Sport team symbol	24
Green text and white background	Peace	21
	Sport team symbol	12
Yellow text and black background	Bugs	32
	Sport team symbol	9
Red text and green background	Christmas	68
	Sport team symbol	6

Table 7.9: The Reflections of Colours' Combinations Meanings on English Users

7.4.3.2.1 Comparing Arabic to English Colours' Findings

The findings confirmed that both Arabic and English users prefer light background colours to the dark. But significant differences were found in the colour combination selection, such that Arabic users preferred dark-blue text over the white background to the black over yellow background, whereas, English users preferred the second combination rather than the first. Also, both the Arabic and English users were attracted by blue background, which was the lightest of all, and both of them rated it as good. However, the findings for the attractiveness of blue showed that Arabic users strongly preferred it as a background. On the other hand, the findings indicated that English users significantly regarded black and pink backgrounds as unattractive.

When comparing the results for arranging the online examples in order of users' preferences, both English and Arabic users identified the blue background as their first choice, and then pink was the second, followed by green and black for the fourth. There was a significant difference with regard to arranging pink and green background examples in the fourth place. Table 7.10 shows these differences.

Regarding the readability measure for the on-line examples, it is noticeable from Table 7.11 that coherent findings were achieved for most of the background colours for both Arabic and English. But there was a significant difference in rating the readability for the example with the pink background colour, such that the English users strongly rated its readability as good with 85%.

Background Colour	Native Language	
	Arabic	English
Blue	0%	0%
Black	57%	68%
Pink	2%	15%
Green	41%	18%

Table 7.10: Background colours Fourth place arranging

Colour	Good		Neutral		Bad	
	Arabic	English	Arabic	English	Arabic	English
Blue	84%	91%	16%	9%	0%	0%
Black	34%	29%	28%	27%	39%	44%
Pink	57%	85%	37%	12%	6%	3%
Green	5%	6%	42%	41%	53%	53%

Table 7.11: Readability measure Findings (for Arabic and English)

It is recommended to use different colour to represent the important text in the web document, since this increased the scanability measure for both English and Arabic users. But the Arabic users preferred this feature in their web pages much more than the English users, and this caused a significant difference between their preferences. There were coherent findings in colour utilisation, since the Arabic and English findings were consistent. Figure 7.1 shows the similarities in the answers.

Despite the divergence between the two cultures, some of the colours' meanings affected both cultures; however the findings showed some cultural preferences. For instance, purity for white and spring for green were highly related in both cultures, but Arabic users significantly supported them more than the English users. On the other hand, envy for green and bright for yellow were closer to English culture than Arabic culture, leaving a significant difference between them.

There are also significant differences in the perception of colour combination meanings between Arabic and English cultures. Arabic users strongly associate the green and white colour combination with peace and red and white with joy; a significant difference from the English users. Alternatively, English users related a red and green colour combination to Christmas, which is an affect of English culture. The realisation of how these differences affected the current usability guidelines is discussed in the next section.

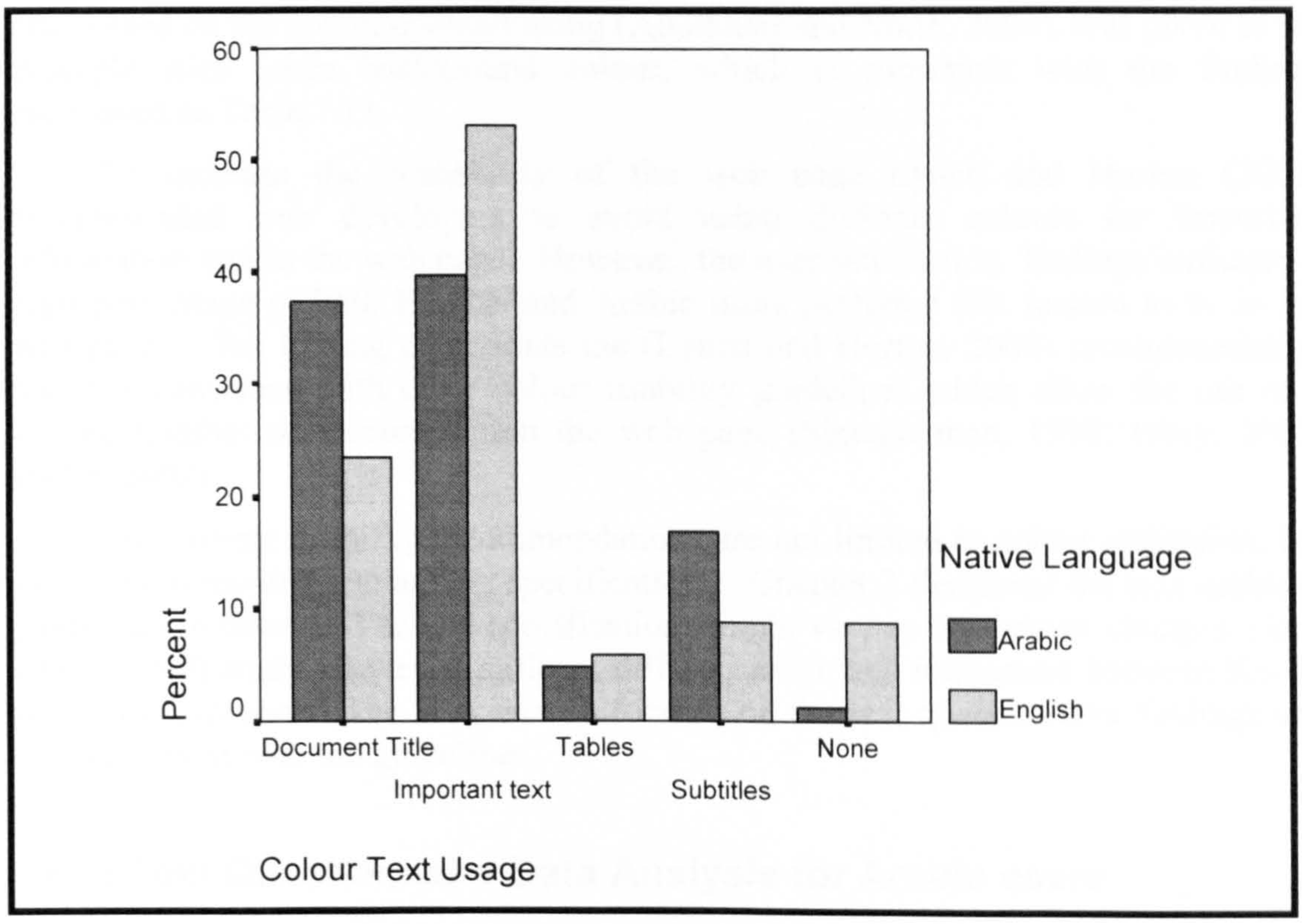


Figure 7.1: Comparing the Colour Text Usage for Arabic and English Findings

7.4.3.2.2 Mapping Between (English and Arabic) Users' Preferences and the Current Colours Usability Guidelines

The details of existing usability guidelines were discussed in Chapter 2. Many of the current web usability guidelines contradict each other. This is partly because the web is so ubiquitous, its users so diverse and their tastes so varied. Usability guidelines appropriate in one culture are not necessarily appropriate in another. Under these circumstances, this section focuses on finding the impact of users' preferences on these guidelines.

Both Arabic and English users preferred the dark text colours over light background colours more than the opposite colour combination. The findings are consistent with those reported by (Lynch and Horton, 2001). Similarly Nielsen and Tahir's (2002) recommendation to select a white background with any dark colour text is consistent with the Arabic users' preferred colour combination. On the other hand, the English users' preferred colour combination was consistent with (Badre, 2002) example of high contrast colour combination whereas, Preece et al,'s (1994) suggestion to use dark background such as blue, with a light text, is inconsistent with both English and Arabic colour combination' preferences.

There must be either high or at least sufficient contrast between text and background colours as (Sklar, 2000; Lynch and Horton, 2001; Badre, 2002; Nielsen and Tahir, 2002) stated for the readability of the web page. The English and Arabic ratings for readability of the four on-line examples were consistent with the previous recommendation. Additionally, these readability ratings were consistent with (Abulkhair and North, 2003), the fuzzy colouring model rating. The lowest usability

rate, based on the colour contrast using (Abulkhair and North, 2003), was given to the example with green background colour, which is consistent with the findings mentioned in Table 7.11.

To increase the scanability of the web page Lynch and Horton (2001) recommended web developers to avoid using different colours for important information within the web page. However, the user satisfaction findings indicated a high percentage of both English and Arabic users preferred this feature to be in the web page. This finding contradicts the (Lynch and Horton, 2001) recommendation but it is consistent with other colour usability guidelines which allow the use of a limited number of colours within the web page (Shneiderman, 1998; Ivory, 2001; Badre, 2002).

The current usability recommendations are not limited to colour utilisation, but extended to involve text and its specifications. Chapter 2 discussed the text usability guidelines in details. The text specifications might vary as the culture changes, since Lee's (1999) study showed significant differences in web interaction between Korea, Japan and America. The next section focuses on the text questionnaire findings and connects them with the guidelines.

7.4.3.3 Text Questionnaire Data Analysis for Arabic users

The Arabic language is an element of the Arabic culture, and Arabic users are affected by their culture. The direction of Arabic writing is from right to left, and the Arabic users web page scanning is clearly influenced by this. The findings showed that 60% of the Arabic users started the scanning from the upper right corner, whereas 18% started from the upper centre. From the different percentages presented in Table 7.12, it can be seen that low percentage rates are given to the other parts of the document.

First point Scanning	Percentage
Upper left corner	5
Upper center	18
Upper right corner	60
Middle of the page	9
None	9

Table 7.12: Arabic users' First Scanning Point

After scanning, the attraction of specific items enclosed in the web page was considered. The web page tested was divided into three sections and each section contained an item like: university logo, document title, questionnaire body. Figure 7.2 illustrates the positions of these items. The findings classified the three items into three levels based on their importance to Arabic users. The university logo was given 49%, document title 33% and questionnaire body 18%. These findings were consistent with the first point scanning findings; so the upper right corner is clearly the most important point in Arabic web pages.

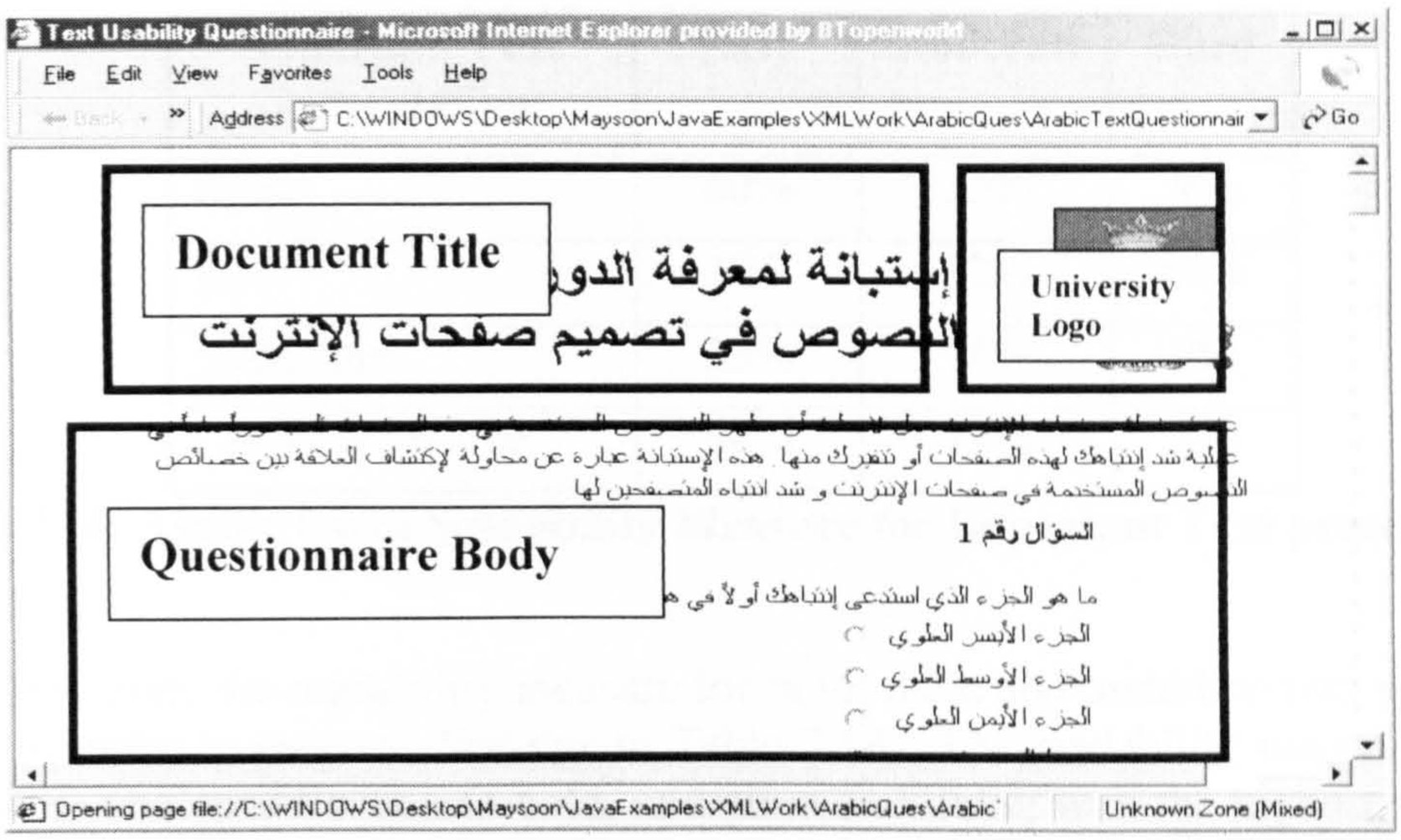


Figure 7.2: The Tested Arabic Web Page contents divisions

In order to emphasize important text, headings and subtitles, the text specification should be tuned to users' preferences. Different coloured text was the preferred way to emphasize important text by Arabic users with 47%, which was similar to the colour questionnaire findings in Table7.3. The second preference was bold with 26%, followed by a larger letter size with 25% and the least popular were underline with 2%, and italic (0%.)

The scanability measure finding for the headings and subtitles was increased by using the following specification: a different colour, bold, larger letter size, italic and underline heading. Table 7.13 involves the demonstration of the scanability measure for each text specification. It is clear that Arabic users strongly preferred coloured heading for fast scanning, followed by bold and then a larger letter size.

Heading Presented in	Percentage
Bold	23 %
Italic	2%
Underline	2%
Larger letter	19%
Different colour	54%

Table 7.13: Arabic users measuring the Heading Scanability

The results from asking the users to measure the scanability of the important text found that a larger letter size was the most commonly preferred presentation for important text. The other presentation preference was a bold text, followed by underline; the least popular was italic. Table 7.14 illustrates the Arabic users' scanability measure for bold, italic, underline and larger letter size.

Important Text Presentation	Easy	Moderate	Hard
Bold	60%	32%	9%
Italic	9%	32%	60%
Underline	33%	58%	9%
Larger letter size	86%	12%	2%

Table 7.14: Arabic Users Scanability Measure for Important Text presentation

However, the readability measure for bold, italic and underline text was similar to the scanability measure findings in Table 7.14. The readability measure findings sorted the text specification to bold, underline and italic, with the highest readability rate identified as pleasant followed by acceptable and then unpleasant. As mentioned earlier bold text presentation was significantly ahead with 70%; and then 56% regarded underline as acceptable, and lastly italic text was seen as unpleasant by 53%. Table 7.15 shows readability rates for the text specifications mentioned in this section.

Readability Rate / Text Specification	Pleasant	Acceptable	Unpleasant
Bold	70%	21%	9%
Italic	9%	39%	53%
Underline	19%	56%	25%

Table 7.15: Arabic Users readability Rates for Some Text specification

Arabic users were subjected to different letter sizes, ranging from 7.5 to 24 point, to identify a precise letter size for users with normal vision. The findings showed that the smallest font size (7.5 point) was strongly rejected by 86% of Arabic users, and considered illegible. The 10 point font was classified as moderately readable by 67% of Arabic users. The rest of the letter sizes tested were considered legible, but the 24 point font size had a lower legibility rate than the 13.5 point. 67% of Arabic users classified the 24 point as legible, but a higher percentage was given to the 13.5 point font size with 91%. It is clear that the maximum legibility measure was given to the 13.5 point letter size and legibility decreases as the font size is either increased or decreased.

The following section will analyse the data for the English version of the text questionnaire.

7.4.3.4 Text Questionnaire Data Analysis for English users

The first scanning point survey showed that 66% of the English users started their scanning from the upper left corner. Based on the scanability measure findings, it is clear that the language writing direction affected the scanability of the web page.

Other parts of the web page had less effects, for example 22% of the English users started their scanning from the middle of the page. The full findings are listed in Table 7.16.

First point Scanning	Percentage
Upper left corner	66%
Upper centre	9%
Upper right corner	3%
Middle of the page	22%
None	0%

Table 7.16: English users' First Scanning Point

Figure 7.3 illustrates the positions of three sections of a web page which were tested to find out the area users are first attracted to. There was no common point that attracted the English users and could be considered as "first look" point. The findings were the same for two sections: the university logo and questionnaire body with 31% each, whereas the document title was given 38%, the highest percentage of all.

The English users considered bold text as the most recognisable presentation for important text, headings and subtitles. The findings indicated that 56% supported bold, whereas the rest of the English users approval was scattered among italic with 9%, underlined with 3%, larger letter and coloured, both with 16%. Capitalised text presentation was not regarded as an attractive way to represent important text by any of the English users.

On the other hand, the scanability measure, for the same six text specifications mentioned (bold, italic, underline, larger letter, capitalised and coloured), was slightly different from the attractiveness measure. The scanability findings for the heading and subtitles arranged in order of importance were bold, larger letter, coloured and capital heading. Table 7.17 shows the exact figure of the percentage dealing with each one.

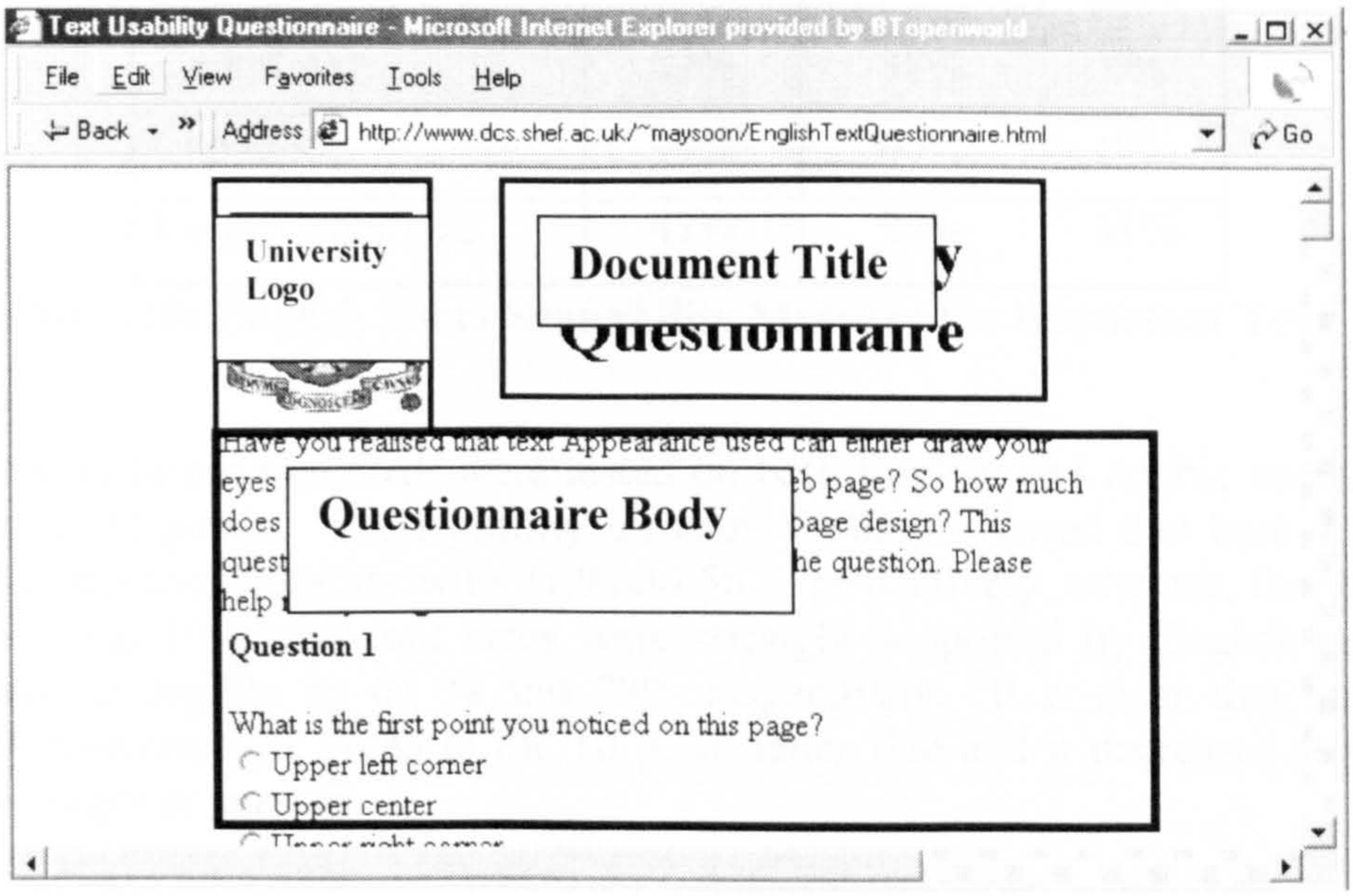


Figure 7.3: The Three Sections of the English Web Page Tested

The scanability measure, for the important text within the web page, showed that the bold text was the fastest scanable text presentation for a large majority of English users. It was considered an easy way to distinguish important text by 78% of these users. The rest of the findings, for the important text scanning, differed from the findings for titles and were, in order of preference, underline, larger letter size, italic and then capital letters. Table 7.18 shows the different percentages given to each of the text specification by associating them with three scanability measures.

Heading Presented in	Percentage
Bold	50%
Larger letter	31%
Capital	3%
Different colour	16%

Table 7.17: English users measuring the Heading Scanability

The readability measure findings differed from the scanability measure for bold, italic, underline and capitalise. The results showed that the readability measure for both underline and capitals were significant in that 81% found them unpleasant to read. However, the readability measures for bold and italic text were considered as acceptable with same percentage 47%. Different percentages for each readability scales are presented in Table 7.19.

Important Text Presentation	Easy	Moderate	Hard
Bold	78%	22%	0%
Italic	25%	53%	22%
Underline	53%	38%	9%
Capitalise	16%	66%	19%
Larger letter size	47%	22%	31%

Table 7.18: English Users Scanability Measure for Important Text

Four different font sizes were tested on both English and Arabic users, ranging from 7.5 to 24 points. The legibility measure findings showed that both 7.5 and 24 points had the low legibility with 71.9 and 56.3, respectively, whereas, the text tested with 13.5 and 10 points font sizes were strongly supported by English users, and considered as legible by 63 % and 78% respectively. It is clear that the highest legibility measure was given to the 10 point letter size and it decreases as the letter size gets larger or smaller.

Readability Rate / Text Specification	Pleasant	Acceptable	Unpleasant
Bold	25%	47%	28%
Italic	34%	47%	19%
Underline	6%	13%	81%
Capitalise	9%	9%	81%

Table 7.19: English Users readability Rates for Some Text specification

The language scripts, as part of the culture, and the users preferences affect the usability measures to some extent. The comparison between the English and Arabic text questionnaire findings is the issue considered in the next section.

7.4.3.4.1 Comparing Arabic with English Text's Findings

It is obvious from these findings that the nature of the language affected the scanability, since the first point scanning differs from one language to another according to its writing direction. This caused a significant difference between Arabic and English in the first point scanning, 60% of Arabic users started their scanning from the upper right corner unlike the English users. About 66% of English users began their scanning from the upper left corner compared to 5% of their Arabic counterparts.

After the scanning process was over, the users started to look at specific item on the web page. Arabic findings appeared to have a more hierarchical structure whereas English findings were in the same level of importance. The top of the web page seemed to be more important to the Arabic users than the rest of the web page; however, their English counterparts appeared to have no significant preference for any part of the web page.

As shown earlier, the majority of English users gave their first attention to bold face for important text presentation, and this was a significant difference from Arabic users. However, the use of different colours present important text within the web pages was considered as the most important text presentation for Arabic users. A high number of Arabic users supported different colour for important text presentation compared with English users, very few of whom preferred that option. Furthermore, the italic text presentation was accepted by a few English users, but it was completely rejected by Arabic users leaving a significant difference between the Arabic and English users' preferences. The comparison between the English and Arabic favourites is illustrated in Figure 7.4.

As a result of the experiment described in earlier chapter, it appears advisable to use at least one of the following text specifications, bold, italic, underline, larger font or different colour in order to differentiate headings and subtitles from the rest of the text. The headings and subtitle scanability measures for both Arabic and English are summarised in Table 7.13 and Table 7.17 respectively. The order of users' favourites differed and Table 7.20 shows the comparison between Arabic and English users' views. The findings showed that bold headings were significantly higher rated by English users than Arabic users who strongly supported coloured headings.

Rating the scanability of important text within a web page gave slightly different results from scanning the headings. By sorting the findings for easy scanning, presented in the second column in Table 7.14 and Table 7.18, it is clear that these findings are fairly consistent with the findings in Table 7.20 in some respects. The consistency with the English findings was more obvious than with the Arabic, such that the first and second choice remain the same as the headings scanability; while the Arabic second choice of the headings findings was in line with the second choice of important text scanning easy measure. Table 7.21 presents the comparison between Arabic and English for important text scanning in order of easy scanning measure.

	Bold	Italic	Underline	Capital	Larger Letter	Different Colour
First	●					▲
Second	▲				●	
Third					▲	●
Fourth		▲	▲	●		

▲ Arabic ● English

Table 7.20: Comparing the Scanability Measure for the Text Specification for Both English and Arabic Users

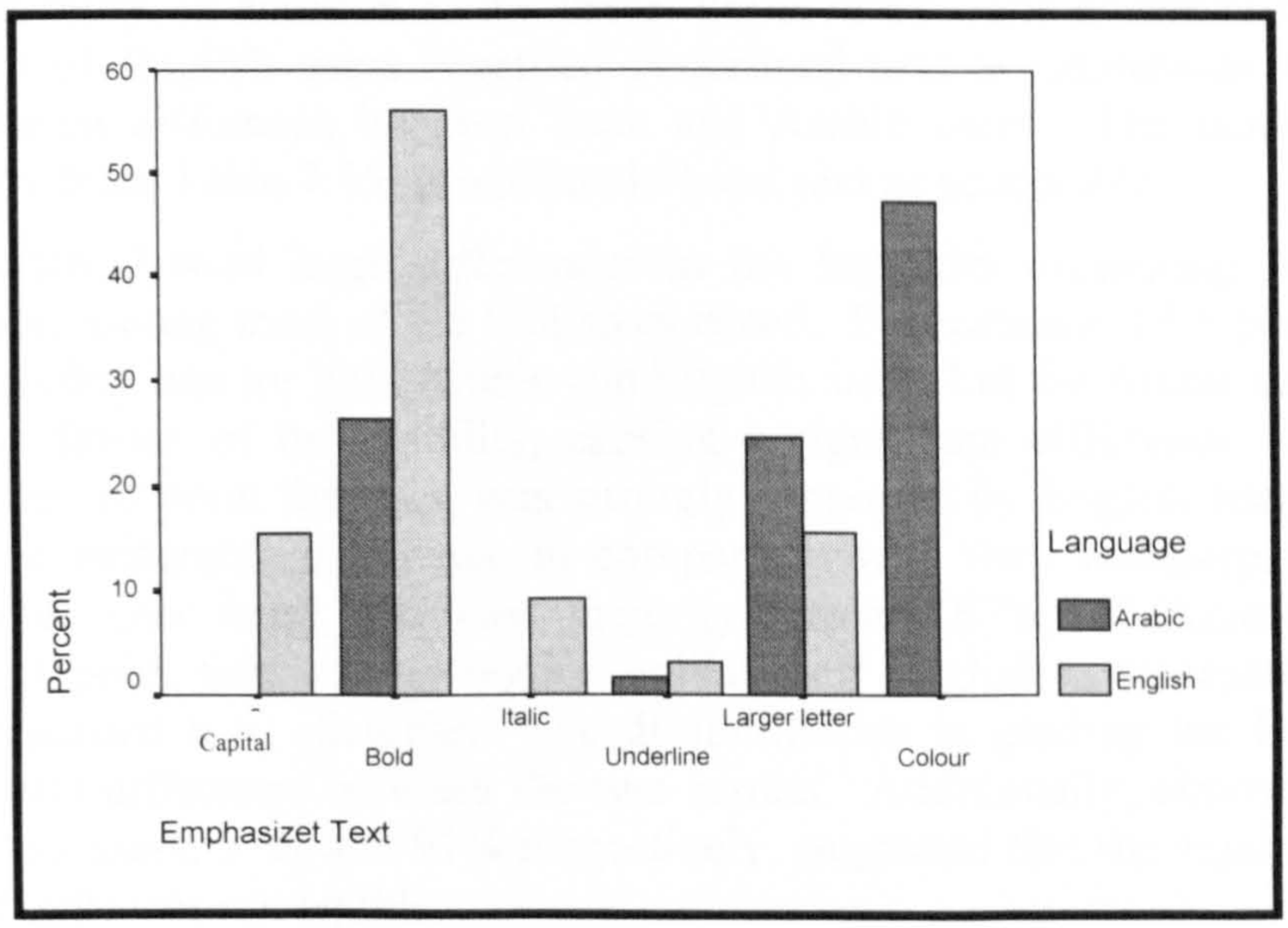


Figure 7.4: Comparing Arabic with English Users Preferences in Presenting the Important Text

	Bold	Italic	Underline	Larger Letter
First	●			▲
Second	▲		●	
Third			▲	●
Fourth		▲		

▲ Arabic ● English

Table 7.21: Comparing the Scanability Measure for Important Text in Order of Easy Scanning for Both English and Arabic Users

Despite the similarities noticeable from Table 7.21 in the third and the fourth order of scanability measures of important text between the Arabic and English findings, there was a significant difference in the italic text presentation, because italic is supported by English users more than Arabic users. However, Arabic users strongly supported larger letters for easy scanning and distinguishing the important text. The high support for larger letter size, from Arabic users, leads to a significant difference between them and their English counterparts.

By comparing the readability measure for the three text specifications (bold, italic, underline), it appeared that large number of Arabic users, see Table 7.15, found it easy to read bold text while few of the English users liked it. Additionally, the majority of Arabic users, see Table 7.15, disliked reading italic text, whereas many (35%) of English users classified it as pleasant to read (see Table 7.19). In both cases a significant difference was observed, which indicates that the language scripts affected the readability, in addition to the users' preferences. The evidence showed that the majority of English users regarded underlined text as unpleasant to read causing a significant difference between them and Arabic users. The majority of Arabic users, 56% from Table 7.15, graded underlined text as acceptable.

This evidence showed huge differences in the legibility measuring between English and Arabic among most of the font sizes tested. For instance, 13.5 point font size had high legibility rate for both Arabic and English users, but the Arabic users are more strongly in favour of its legibility, causing a significant difference with the English. However, 10 point font size was strongly supported by English users only, again leading a considerable difference in comparison with their counterparts (the Arabic). On the other hand, the vast majority, around 67%, of Arabic users considered the 24 point font size as legible unlike their English counterparts, who mainly (56%) regarded it as illegible. The dissimilarities in grading the legibility caused a significant difference between the two results. Additionally, almost all the Arabic and English users, 97% and 91% respectively, suggested that the smallest font size tested (7.5 point) was not legible.

From all the comparisons of findings discussed earlier, there were significant differences in some of the text features. These differences affected the usability guidelines, which should be localised to match the users' preferences. The next

section discusses the compatibility and the differences between the findings and the current web usability guidelines.

7.4.3.4.2 Mapping Between (English and Arabic) Users' Preferences and the Current Text Usability Guidelines

By recalling the usability guidelines, discussed in Chapter 2, it is clear that those guidelines are appropriate for some groups of users but not suitable for others. For instance, the basic guideline for the web page traversing described by (Sklar, 2000) seemed to be based on English, but might be appropriate for any language with a Latin script, since the majority of English users, more than 65% (see Table 7.16), were consistent with Sklar's traversing guideline. However, this guideline is not suitable for Arabic script, because only a very small number of Arabic users, around 5% presented in Table 7.12, started their traversing from upper left corner.

Sklar (2000) divided the web page into 5 sections arranged in order of importance demonstrated in Figure 7.5. The findings for both Arabic and English were inconsistent with the Sklar's depiction (2000). Arabic findings showed that almost half of the Arabic users, 49%, first look at the university logo which was on the number 2 of the Sklar's important areas (2000), whereas the English findings, for the first point looked at in the web page, were incoherent. In other words there were no significant favourites; so this is also inconsistent with the Sklar's suggestion (2000).

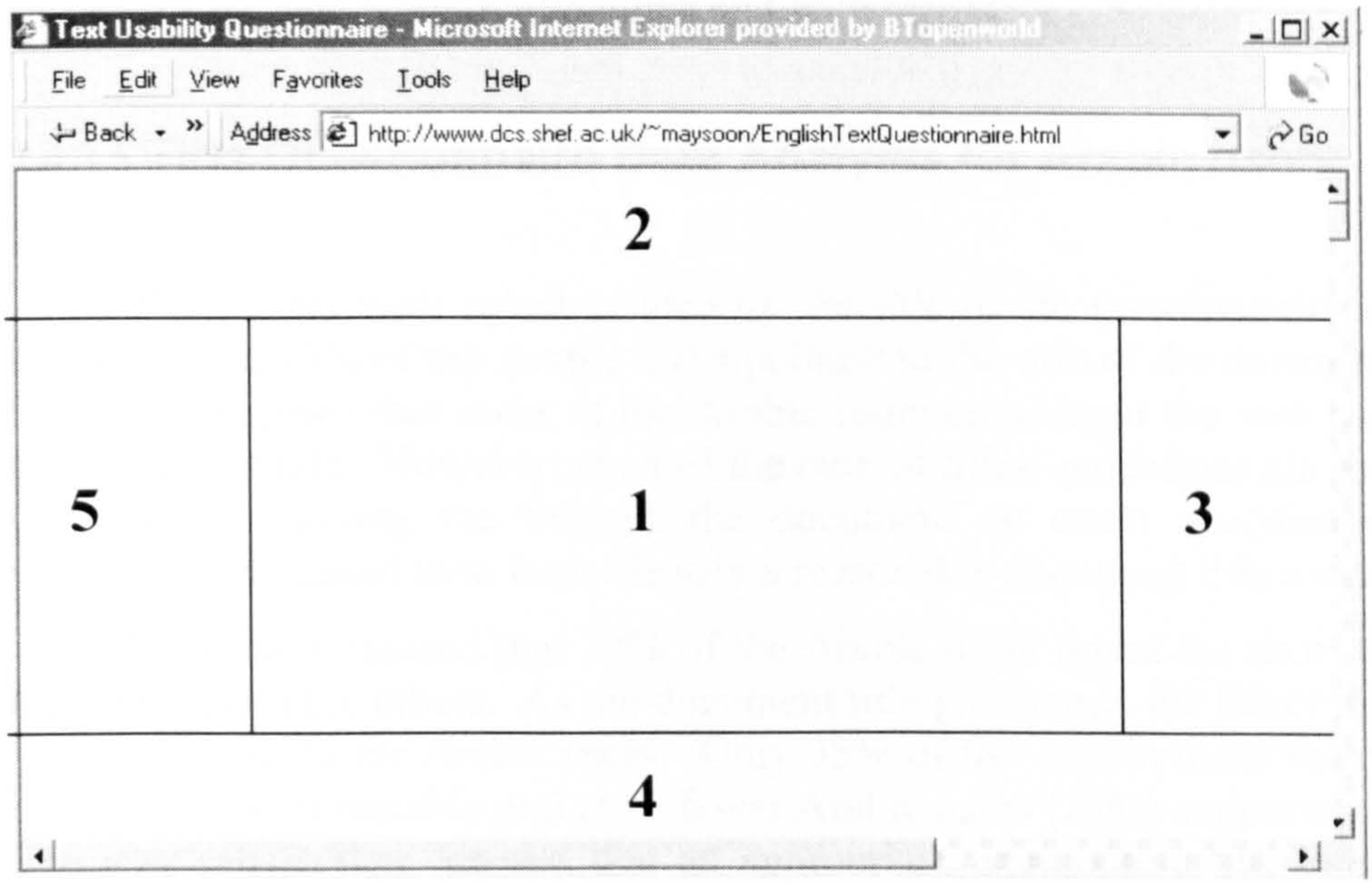


Figure 7.5: Sklar's Relative Areas of Screen Importance

Arabic users preferred a different colour to emphasize important text, headings and subtitles, which contradicts (Lynch and Horton, 2001). However, Lynch and Horton's (2001) recommendation to use bold face was consistent with English users findings since the majority of English users preferred the bold face for important text, headings and subtitles presentation.

Around 53% of the Arabic users (see Table 7.15) found it difficult to read large blocks of italic text, which is consistent with (Sklar, 2000; Lynch and Horton, 2001). However, the English users gave the italic the highest reading rate among the other text specification and they considered it as acceptable to read. This contradicts (Sklar, 2000; Ivory, 2001; Lynch and Horton, 2001), whereas the English findings for the underline readability were in line with (Ivory, 2001; Lynch and Horton, 2001). But less than 25% of the Arabic users thought underlining unpleasant, which was inconsistent with the literature (Ivory, 2001; Lynch and Horton, 2001). The majority of English users found that it was hard to read large blocks of capital letters and this was consistent with (Nielsen, 2000; Ivory, 2001). The English users did not like the underlined text even in headings, subtitles or for important text, which is inconsistent with (Lynch and Horton, 2001; Badre, 2002; Nielsen and Tahir, 2002).

According to the font size literature, (Ivory, 2001; Lynch and Horton, 2001) recommend 14 point or higher for headings, which is consistent with the Arabic users preferences. The Ivory (2001) font size recommendation is limited the text font size from 10 to 11 points, this recommendation is consistent with the English findings, because they gave the highest legibility rate to such size. However, for the vast majority of Arabic users, the maximum legibility rate was given to the 13.5 point font size, which was inconsistent with (Ivory, 2001; Lynch and Horton, 2001).

It is clear that both Arabic and English findings for text specification have some consistency with the current web usability guidelines; on the other hand they conflicted on some occasions. The following sections discuss the analysis of the third part of the questionnaire, which was aimed at validating some other web specifications dealing with title, links and alignments.

7.4.3.5 Title Questionnaire Data Analysis for Arabic users

When users were asked to identify the title of the questionnaire web page the findings were 73% of the Arabic users pointed to the title of the document. From the findings it appears that most of the Arabic users considered the web page title to be the document title. However, most of the current titles' guidelines are for the browser titles without giving the title of the document an exact specification. Further questions were asked to at least identify a reasonable document title length.

The results showed that 78% of the Arabic users found the short titles easier to recognise them the others. As the document title got longer, the lower the chance of it being memorable for Arabic users. Only 35% of the Arabic users rated the medium title length as memorable and even fewer Arabic users (22%) supported the long title. The user satisfaction proved that an appropriate, memorable title length should be within 20 characters in length and as the title length increased the memorability decreased. Many Arabic users did not like many links in a web page. The findings showed that 62% of them dislike to have several links.

Since Arabic writing is from the right to left, the heading alignments tested were based on right alignment for text body. Headings alignment findings demonstrated that right alignment for headings was considered as the most organised alignment of all. Centred headings came second, and left heading alignment was regarded as disorganised by the vast majority of Arabic users. Table 7.22 illustrates how Arabic users regarded heading alignments.

Heading Alignment	Organised	Moderate	Unorganised
Right	71%	27%	2%
Centre	55%	44%	2%
Left	6%	11%	84%

Table7. 22: Arabic Users Rates for Different Heading Alignments

The findings of the readability measure for the text body alignment showed that 75% of Arabic users rated right text alignment as easy to read. Centre text alignments was graded as moderate, followed by left text alignment which was rated as hard to read. These findings were consistent with the Arabic writing direction.

7.4.3.6 Title Questionnaire Data Analysis for English users

The findings showed that 75% of the English users regarded the title of the document to be the title of the web page. But current guidelines are aimed at the browser's title and include the title of the document with the headings. English users were involved in evaluating the length of different documents titles.

The English users' evaluation results showed that 71% of them preferred a short document title. However, 61% of the English users rated the medium length title as moderately memorable. A long title was rated as forgettable by 61% of the users. It was obvious that each time the document title became longer, it increased the probability of its contents being forgotten. The findings indicated that English users preferred many links in the web pages they browse, since 54% of them like to see several links in the web page.

The results for the most preferable heading alignment was the left alignment. The English users found that web pages examples, with a left heading alignment, were more organised than the others. The findings showed that the web page organisation rate lowered, as the heading alignments moved to the centre or to the right. Table 7.23 shows the evaluation rates for different heading alignments. It is clear that the left heading alignment was seen as the most organised web page by English users, which is consistent with the English writing direction.

Heading Alignment	Organised	Moderate	Unorganised
Right	11%	43%	47%
Centre	50%	47%	4%
Left	79%	18%	4%

Table7. 23: English Users Rates for Different Heading Alignments

7.4.3.6.1 Comparing Arabic to English Title Findings

It is obvious that there were some similarities in the title questionnaire findings between English and Arabic. For instance, both Arabic and English users considered the document title to be the web page title instead of the browser title, with close percentages of 73% and 75% respectively. Also, both English and Arabic findings were consistent in determining the most memorable document title length. Each of them believed that the short title length, which was around 20 characters long, was more memorable than any others. Thus, when the title length increases to more than 20 characters, normal users' memories began to lose concentration.

However, there was a difference, on the subject of including many links in the web page, between English and Arabic users. More than half of the English users preferred to have many links in the web pages unlike their Arabic counterparts. Around 62% of Arabic users disliked many links in the web page.

There was a significant difference between Arabic and English right headings alignments evaluation. Arabic users strongly supported the right headings alignments because of the Arabic language writing direction; so a dramatic but predictable difference appeared between Arabic and English in rating left heading alignments. The English users preferred left headings to any other (centre and right), consistent with the English writing direction. However, the findings of the centre headings alignments were close to each other; in other words, there was no significant distinctions between the Arabic and English results. Figure 7.6 compares the Arabic and English centre headings alignments findings.

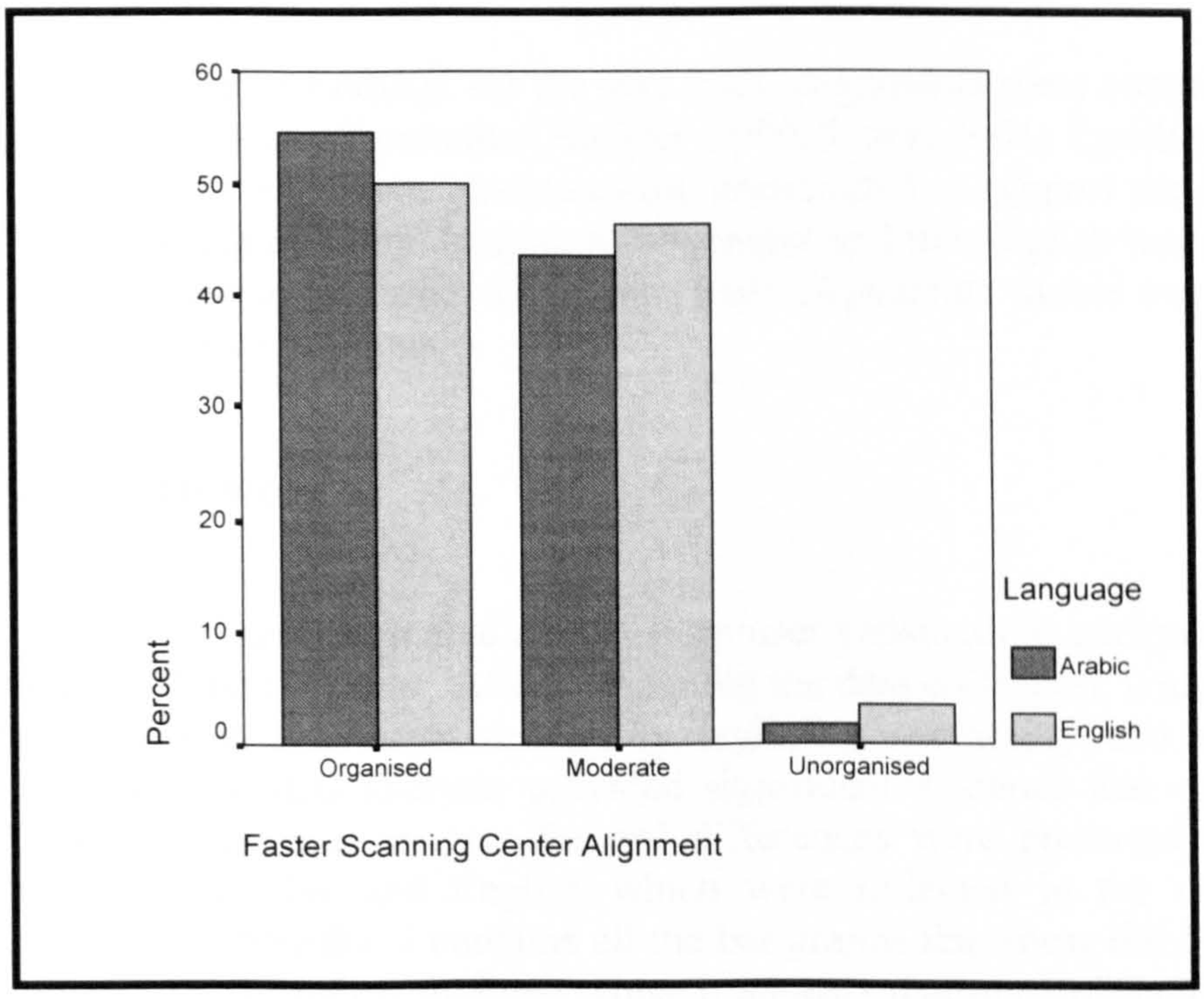


Figure 7.6: Comparison Between the Arabic and English Centre Heading Alignments

7.4.3.6.2 Mapping Between (English and Arabic) Users' Preferences and the Current Title Usability Guidelines

Most of the current usability guidelines consider the browsers' title as the web page title (Nielsen, 2000; Ivory, 2001; Nielsen and Tahir, 2002), and have established different recommendations to characterise it. The title characteristics were discussed in Chapter 2. However, most of the users tested, for both Arabic and English, considered the document title as the web page title. Users from dissimilar communities thought that the text that appeared at the top of the web page was the web page title, and not the title located on the top of the browser's window frame.

Nielsen's recommendation (2000) for minimising the number of the links within a page was consistent with the Arabic findings; a high number of Arabic users did not like many links within the web page. But the English findings contradicted the (Nielsen, 2000) guideline, since the majority of the English users preferred to have a large number of links.

Lynch and Horton's (2001) recommendations for headings alignment were consistent with the English findings. English users significantly supported the left headings alignment on the basis of the left aligned text, unlike their Arabic counterparts. On the other hand, the right headings alignments were inconsistent with the (Lynch and Horton, 2001) suggestions that claimed that centre and right headings alignments make the display unbalanced. The right headings alignments (Lynch and Horton, 2001) recommendation were in line with the English findings. Furthermore, the centre headings alignments (Nielsen, 2000) guidelines were consistent with both findings, since Arabic and English users described this kind of headings alignments as moderate.

The Arabic findings for the text body alignments were completely inconsistent with the literature alignments (Nielsen, 2000; Ivory, 2001; Lynch and Horton, 2001) as most of the alignment guidelines are constructed to support the Latin script or left to right writing direction languages in general and the English language in particular. Thus, the measuring scales for the text body alignments should not be the same if the language direction changed.

7.5 Conclusion

This chapter presented results from user satisfaction performed by 290 Arabic and English participants. After completing the data collection, it analysed this data on the basis of attractiveness, scanability, legibility, readability, and aspects' utilisation. The subjective data analysis provided significant evidence that culture has a major effect on web page design. Several differences were produced by comparing the findings of Arabic and English which were reflected in the usability guidelines validation. Appendix *A* contains all the bar graphs that show the differences between the Arabic and English findings. Thus, it appears that most of the usability guidelines were established for English, but still need to be updated frequently to meet the users' changing views and technological developments.

Since culture can affect the way the users perceive the web pages, the usability guidelines need to be localised. In order to build a reliable, automated usability

evaluation tool, it should be possible to customise its evaluation to the users' needs. Additionally, the automated evaluation tool should be prepared to update its evaluation rules to be consistent with changes in the users' preferences. The usability rules adopted were captured from both the user satisfaction results and from the current usability guidelines. The basic usability rules are considered in the evaluating agent's construction and will be examined over several Arabic and English examples. The agent's evaluation is discussed in the next chapter.

Chapter Eight

Assessing the Agent's Evaluation

8.1 Introduction

Chapter 7 linked the user satisfaction, with the current usability guidelines and pointed to the significant differences between the Arabic and English users' preferences. The results from Chapter 7 are adopted in the fuzzy evaluation model, and they form the basic usability rules of the inference system of the model.

However, this chapter discusses the use of the fuzzy model implemented in the evaluation phase of the agent depicted in Chapter 6 Section 6.5.2. The intent of this chapter is threefold:

- Test the effectiveness of the fuzzy model in producing reasonable evaluation in the light of the guideline and the usability testing findings.
- Show the applicability to use the agent to measure the readability of dynamically changing web pages.
- Assess different types of web pages. In other words, the readability of different web pages' categories can be evaluated by the agent.

As mentioned in Chapter 6 Section 6.5, the multilingual automated web usability evaluation agent is built on the four different phases, of analysis, evaluation, learning and controller. This chapter presents the outcomes of the first two phases. Firstly, it discusses the quantification of the web page content performed in the analysis phase of the agent. Secondly, it applies the different web metrics quantities computed to produce the usability rate. Thirdly, it decides if the fuzzy model fulfils its targeted goals. The agent performance description is associated with two examples, one for each language included in the study and described in the next section.

8.2 The Web Pages Examples

To illustrate the function of the multilingual automated web usability evaluation agent in both English and Arabic environments the agent has been applied to both the BBC and the Aljazeera homepages, as samples of English and Arabic web pages respectively. The BBC and Aljazeera are well-known broadcasting organisations and their web pages have a reliable informational reputation, so they attract a wide audience. Additionally, as news providers, the BBC and Aljazeera are both fast changing web sites. Figure 8.1 and Figure 8.2 show the layout of both web page examples. Both web pages provide a wide range of information which involves news, business and money, health, sports, etc.. The English web page example was selected because, at first sight the BBC web page appears to have a good usability features

such as: high colour contrast, suitable font size, small and few images and a consistent layout. It is noticeable that the Aljazeera web page has different usability features well-matched with the Arabic users preferences discovered in Chapter 7 Section 7.4.3.1; for instance in the contrast between the background and the text colour, the font size and the use of different colours for presenting the subtitles.

The process of evaluation starts when the agent receives the URL address of the targeted web page. In other words, when the agent receives the BBC or Aljazeera URL address, it downloads the HTML source code and starts the analysis.

Figure 8.1: BBC Homepage (http://www.bbc.co.uk/) October 2003

8.3 Web Page Examples Analysis

As described in Chapter 6 Section 6.5.1, the analysis phase of the agent would perform the HTML tags identification and the web metrics computation. This step is necessary to quantify the web aspects involved in each web page and investigated in the study (see Chapter 6). The current research used a total of fifteen web metrics to measure the usability of both Arabic and English web pages. These web metrics were considered as basic factors in the production of a legible web page. The computation of each web metrics depends on its availability in a web page, for instance if the web page did not include an italic font effect, then the analysis would be zero. The following discussion contains the analysis of both Aljazeera and BBC web pages.

Figure 8.2: Aljazeera Homepage (http://www.aljazeera.net/) October 2003

8.3.1 Aljazeera Web Page Analysis

Figure 8.2 shows that the Aljazeera web page contained many images. The metric computation showed that 78 images were included in Aljazeera web page and the summation of their sizes was 3148983 bytes. Aljazeera web page findings also contained 1436 words. The values for both italic and underlined text were zero. The collected data showed that 79 links were used in Aljazeera web page. The second column in Table 8.1 shows the web metrics collected from the Aljazeera Arabic web page.

8.3.2 BBC Web Page Analysis

It is noticeable from Figure 8.1 that fewer images were used in the BBC web page. The analysis results showed that 35 images were included in the design of the BBC web page with 30185 bytes summation sizes. The collected data revealed the total number of words used was 464. There were 213 links registered in the BBC web page. Both the underlined and italic text effects were recorded as zero, since they had not been used in the design. Other collected data can be found in the third column of Table 8.1.

As can be seen from Table 8.1 there is a difference between the analysis of both Aljazeera and BBC web pages, which may relate to the users' preferences mentioned

in Chapter 7. The agent's evaluation of the BBC and Aljazeera is different because it is based on the users' preferences and is discussed in the next sections.

8.4 Comparing Aljazeera with BBC Web Pages Analysis

In spite of the calls for internationalising the web pages (Russo and Boor 1993), the current usability guidelines seem to pay little attention to cultural differences that are reflected in user preferences. Applying current usability guidelines, aimed to assess English web pages, to evaluate web pages implemented in other languages such as Arabic is not necessarily reasonable, since these guidelines reflect the users' preferences in a particular society, and embed the users' cultural values as well as more practical differences such as writing/reading direction.

It is quite clear from Table 8.1 that there were some differences, between the BBC and Aljazeera web pages, at least in the web aspects investigated in the study. For instance, Aljazeera's web page size and the number words contained in it were much larger than the BBC's. The evidence also showed that number of links in the BBC web page was proportionally greater than its counterpart by a factor of three, since Aljazeera web page had 79 links whereas the BBC contained 213. However, the number of images in Aljazeera web page was pro rata more than double that in the BBC page. It was noticeable that the BBC browser title length was three times as long as the Aljazeera browser title. There is only one font size analysed in Aljazeera web page and it was size 12 point, but the BBC's font sizes varied from 12 to 7.5 point.

All of these differences were consistent with the usability testing conducted by this research (see Chapter 7) and provide an informal confirmation of them. Therefore, the evaluation for each web page was based on the users' (Arabic or English) preferences established by the usability testing in addition to the current usability guidelines.

Web Metric	Aljazeera Values	BBC Values
Web page size	41125	37188
Word count	1436	464
Bold words	0	69
Italic words	0	0
Underlined words	0	0
Number of links	79	213
Heading alignment	0	0
Text alignment	3, {right, centre, left}	3, {right, centre, left}
Number of images	78	35
Images sizes	3148983	30185
Words in browser title	4	11
Font sizes count	1	3
Font size	12	12, 10, 7.5

Table 8.1: Aljazeera and BBC web pages metrics computed by the analysis phase of the agent

8.5 Web Page Examples Evaluation

The evaluation of the web metrics was derived from the current usability guidelines in combination with the user satisfaction conducted by this research (see Chapter 7). Additionally, the web usability evaluation was based on the fuzzy evaluation model, discussed in Chapter 6, Section 6.5.2, and implemented in the evaluation phase of the agent.

As mentioned in Chapter 6, Section 6.5.2, the evaluation of the web metrics collected from the analysis phase is classified and tested by an appropriate fuzzy model. These fuzzy models varied to test the length, contrast, size and the repetition of different web metrics such as links, bold, italic and underlined font effects. The following two sections will present the findings of the agent's evaluation for Aljazeera and BBC.

8.5.1 Aljazeera Web Page Evaluation

Table 8.2 shows the Agent's primary usability evaluation findings for the Aljazeera web metrics collected. The evaluation findings for both the font size used and the bold text effect in Aljazeera web page was 100. There were many images included in the design as illustrated in Table 8.1, and then their usability evaluation result was zero. Seventy-nine links were used in the web page design, therefore the links count evaluation result was 33. The rest of the web aspects usability evaluation is revealed in Table 8.2.

Web Aspect	Fuzzy Evaluation
Font size = 12	100
Links count	33
Browser title word count	98
Right text alignment	100
Centre text alignment	50
Left text alignment	10

Table 8.2: The Agent Primary Usability Evaluation of Aljazeera Web Page

8.5.2 BBC Web Page Evaluation

The usability evaluation results for the registered web aspects available in the BBC web page were: 91 for bold effect, zero for the image sizes since there were 35 images (see Table 8.1). The findings for the font size were 100, 91 and 65 based on three, two and one font sizes respectively.

Web Aspect	Fuzzy Evaluation
Font size = 12	100
Font size = 7.5	65
Font size = 10	91
Bold text effect	91
Links count	100
Browser title word count	59
Left text alignment	100
Right text alignment	10
Centre text alignment	50

Table 8.3: The Agent Primary Usability Evaluation of the BBC Web Page

The links count evaluation was 100, since there were 213 links within the BBC web page and as mentioned in Chapter 7, Section 7.4.3.6 the English users prefer to have many links in their web page. Table 8.3 summaries all the usability evaluation produced by the agent.

8.6 Comparing the Evaluation with the Existing Usability Guidelines

The font size evaluation was consistent with both the usability testing (see Chapter 7, Section 7.4.3.3 and Section 7.4.3.4) in combination with the current usability guidelines (Nielsen 2000; Ivory 2001; Lynch and Horton 2001; Nielsen and Tahir 2002), since many usability recommendations suggested that bold effect text should be minimised within the same web page (Lynch and Horton 2001). The findings showed that the BBC bold percentage was 14% and was given 91 usability rates. If the bold usage increased the usability rate would decrease. One of the (Nielsen 2000) recommendations was to minimise the number of images used in the design, because they affect the web page download time. The findings showed that Aljazeera and the BBC contained many images, with large summation sizes illustrated in Table 8.1, and their usability rate were zero. The BBC link count usability rate was 100 and that was compatible with the usability testing findings discussed in Chapter 7, Section 7.4.3.6. However, Aljazeera links count usability rate was 33, much lower than its counterpart because the Arabic users do not like to have many links within the web page.

As mentioned in Chapter 2, Section 2.3, most of the current guidelines recommended using a maximum of six words for the title (Nielsen 2000; Ivory 2001) and together with the findings from Chapter 7, Section 7.4.3.5, Aljazeera browser title was given a high score (see Table 8.2). Other browser's title usability guidelines recommended using no more than eight words (Nielsen and Tahir 2002). Thus, the evaluation of the BBC browser's title was reduced to 59 since it contained 11 words. Even though there are many recommendations suggested to minimise the number of colours used in the web page (see Chapter 2), the findings from the usability testing showed that Arabic users prefer to have many colours (see Chapter 7). The text alignment findings were consistent with the usability guidelines mentioned in Chapter 2 together with the Chapter results.

8.7 Comparing the Agent's Evaluation with the Existing Usability Evaluation Tools

Even though there exists a variety of usability evaluation tools, mentioned in Chapter 3, almost all of them have dissimilar assessment interpretation. Thus, the evaluation output produced by the agent cannot be compared with any of the current evaluation tools for two reasons. Firstly, this agent has the ability to evaluate Arabic web pages which the current usability evaluation tools cannot do. Secondly, even when evaluating English web pages, existing usability evaluation tools introduce their results differently, and no means of direct comparison exist.

8.8 Dynamic Changes in the Web Pages

One of the advantages of automated web usability evaluation tools is that they can assess the web pages at any time. The web environment is dynamic since different web pages are frequently updated, therefore to ensure the quality of the improved web pages, they need usually to be evaluated. Thus, it is necessary to automate web usability assessment so that changes can be detected and the new pages evaluated in line with the webmaster's requirements.

This section focuses on the dynamic changes in web pages, especially on the news broadcasting examples presented in the previous section. By testing the BBC and Aljazeera web pages through selecting them at random times, it can appear that different web aspects were changed and these changes will be reflected in the scores for the usability. Table 8.4 and Table 8.5 show the changes in the Aljazeera and BBC web pages over time.

There were differences among the analysis in various times for both the BBC and Aljazeera. These differences are in the words from the web page, words count in bold, number of links, number of images, image sizes in bytes and the web page size in bytes. The modified web aspects affect the usability scores produced by the agent. The next section discusses the influence of different web metrics quantities analysed in determining the usability rate.

8.9 Assessing Different Web Pages Categories

The agent's evaluation is not restricted to the evaluation of a specific category of web pages such as the news broadcasting. Since the principles for measuring readability as an aspect of web usability is not restricted to any particular type of web page, the usability guidelines adopted can be applied to other categories; however the presentation might differ. In other words, readability is independent of the content or subject area. This is demonstrated below. The next section describes the agent's ability to measure readability as an aspect of web usability for different web categories.

Web Aspect	Aljazeera Web Metrics				
	Aug. 14	Sep. 3	Sep. 8	Sep. 9	Oct. 14
Web page size	35068	42446	43238	44096	41125
Word count	1343	1434	1439	1490	1436
Bold words	0	53	0	35	0
Italic words	0	0	0	0	0
Underlined words	0	0	0	0	0
Number of links	55	76	82	82	79
Browser title length	4	4	4	4	4
Text alignments	3, {right, centre, left}	3, {right, centre, left}	3, {right, centre, left}	3, {right, centre, left}	3, {right, centre, left}
Number of images	39	76	82	85	78
Image size	1310091	3225896	3545516	3748160	3148983
Font sizes count	1	1	1	1	1
Font size	12	12	12	12	12

Table 8.4: Web Metrics Collected from Aljazeera Web Page in Several Times

Web Aspect	BBC Web Metrics				
	Aug. 14	Sep. 3	Sep. 8	Sep. 9	Oct. 14
Web page size	36639	36914	36911	37051	37188
Word count	462	467	465	467	464
Bold words	65	67	70	64	69
Italic words	0	0	0	0	0
Underlined words	0	0	0	0	0
Number of links	212	212	213	213	213
Browser title length	11	11	11	11	11
Text alignments	3, {right, centre, left}	3, {right, centre, left}	3, {right, centre, left}	3, {right, centre, left}	3, {right, centre, left}
Number of images	34	35	35	35	35
Image size	23085	26881	29693	26470	30185
Font sizes count	3	3	3	3	3
Font size	{7.5,10, 12}	{7.5,10, 12}	{7.5,10, 12}	{7.5,10, 12}	{7.5,10, 12}

Table 8.5: Web Metrics Collected from the BBC Web Page in Several Times

Four different web pages, for both Arabic and English (see Figure 8.4), from dissimilar web categories have been assessed; banking, charity, government, and family web pages. These examples were chosen because they targeted different groups of users and exhibited some dissimilar web aspects such as the colours used, and font sizes. Alahli Bank and Lloyds Bank are examples of the banking category in Arabic and English respectively, while for charity and government web pages Red Crescent and Saudi Post and the Red Cross and Royal Mail were chosen.

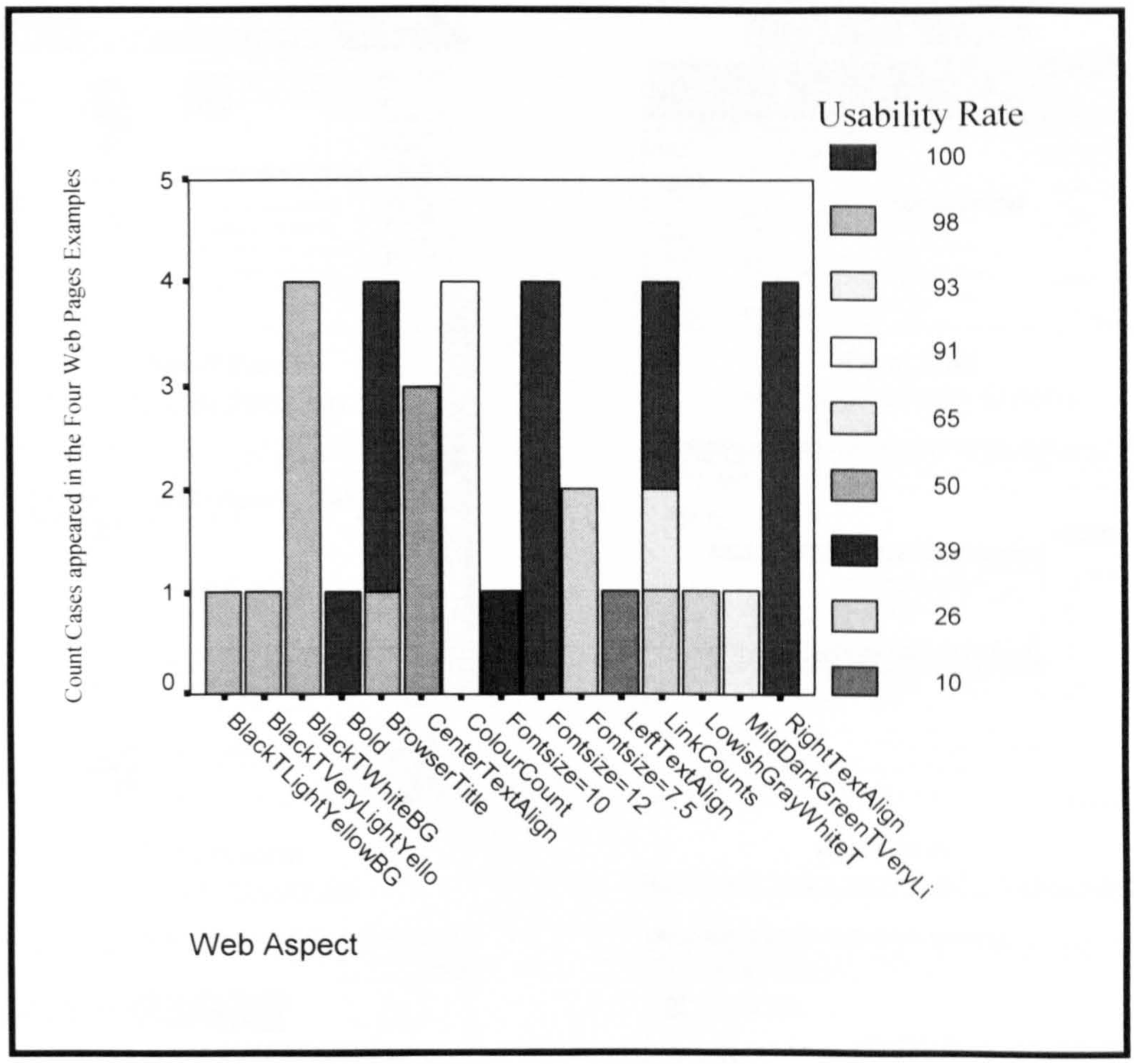


Figure 8.3: The Evaluation Analysis for Different Arabic Web Pages Examples

The evaluation of different web aspects, appearing on the right side of Figure 8.3, ranges between 10 to 100. Each bar in the figure represents one aspect in the four Arabic sample web pages. The greater variation appears in the "number of the links" category, since they vary from one web page to another and their values are 26, 65 and 100. It is clear from Figure 8.3 that two different usability rates are assigned to the evaluation of the browser title length. The usability rate for the rest of the web aspects are either the same in all the web pages tested, appear with the same colour, or the web aspect appeared individually and was not repeated.

Arabic Examples

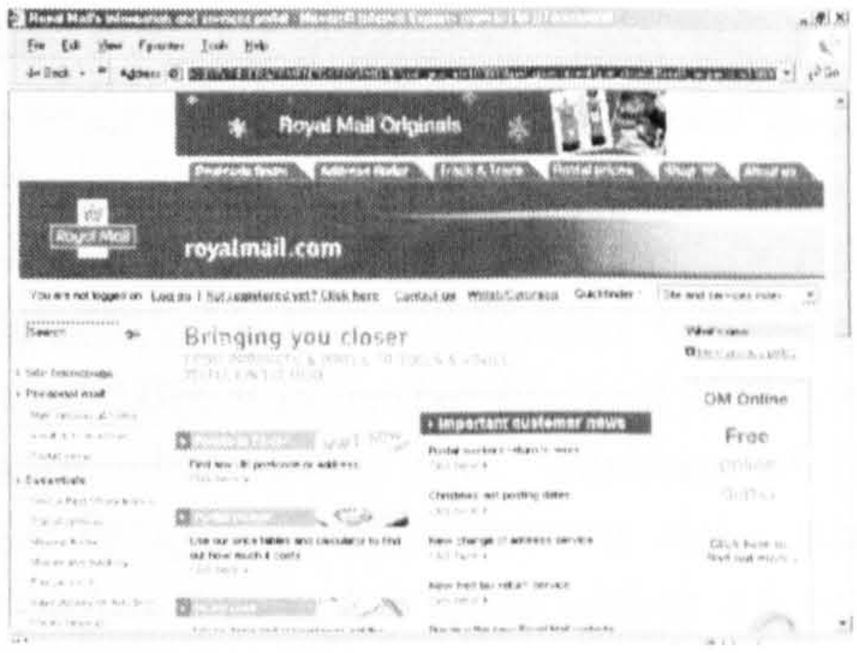
Saudi Post
http://www.saudipost.gov.sa/

Red Crescent
http://www.srcs.org.sa/

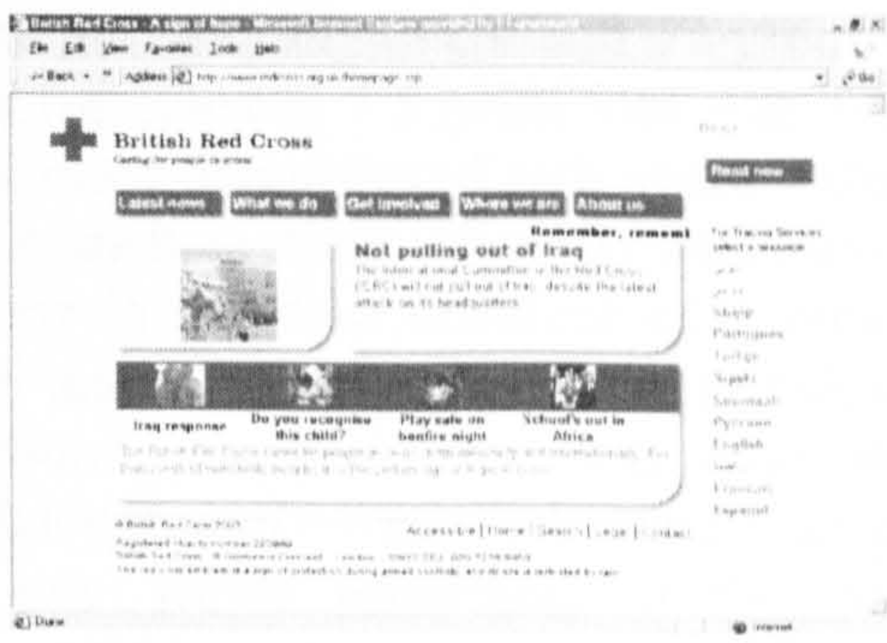
Alahli Bank
http://www.alahli.com/arabic/

Laha Online
http://www.lahaonline.com/People/index.htm

English Examples



Royal Mail
http://www.royalmail.com/



Red Cross
http://www.redcross.org.uk/homepage.asp

Lloyds TSB
http://www.lloydstsb.com/

Families Online
http://www.familiesonline.co.uk/main/topics.htm

Figure 8.4: (Government, Charity, Families, Banking) Arabic and English Examples Taken in November 5, 2003

The variety of the web page aspects analysed and evaluated from the English web pages are represented in Figure 8.5. Even where some aspects appeared in all the four example web pages there are some differences in the usability rate based on their quantities in the web page. From Figure 8.5 it is apparent that the usability rate for both the "bold" and "number of links" web aspects take three different usability rate values. The bold usability rate ranges 65, 91 and 100, where the highest score based on very few is bold utilisation in the web page. However, the "number of links" usability rate varies between 39 and 91. The length of the browser title and the number of the colours utilised in the four web pages examples were either 85 or 100, also based on their representation in each web page.

It is obvious that the other bars in the figure seem to have the same usability rate in all the cases (the four examples) such as the left text alignment and 12 point font size. The usability rate might be the same in three cases as 7.5 point font size, which is size one, and in right text alignment. Several other quantities of web aspects are the same in two cases or appeared in only one case such as italic where its usability rate was 65. Furthermore, the colour combinations in the four different examples are black text over white, highish green, light grey and very light grey background, dark brown text over light green, white background, medium blue text over white background, medium grey text over white background, very dark blue text over darkish red background. Their usability rate varies according to the colour combination contrast between 26 and 98.

It is clear from all the usability scores among the examples assessed that usability rates were consistent if the quantities of different web aspects are equal. But when different quantities of specific web aspect used in several web pages, it causes an individual usability rate, which causes differences in the evaluation.

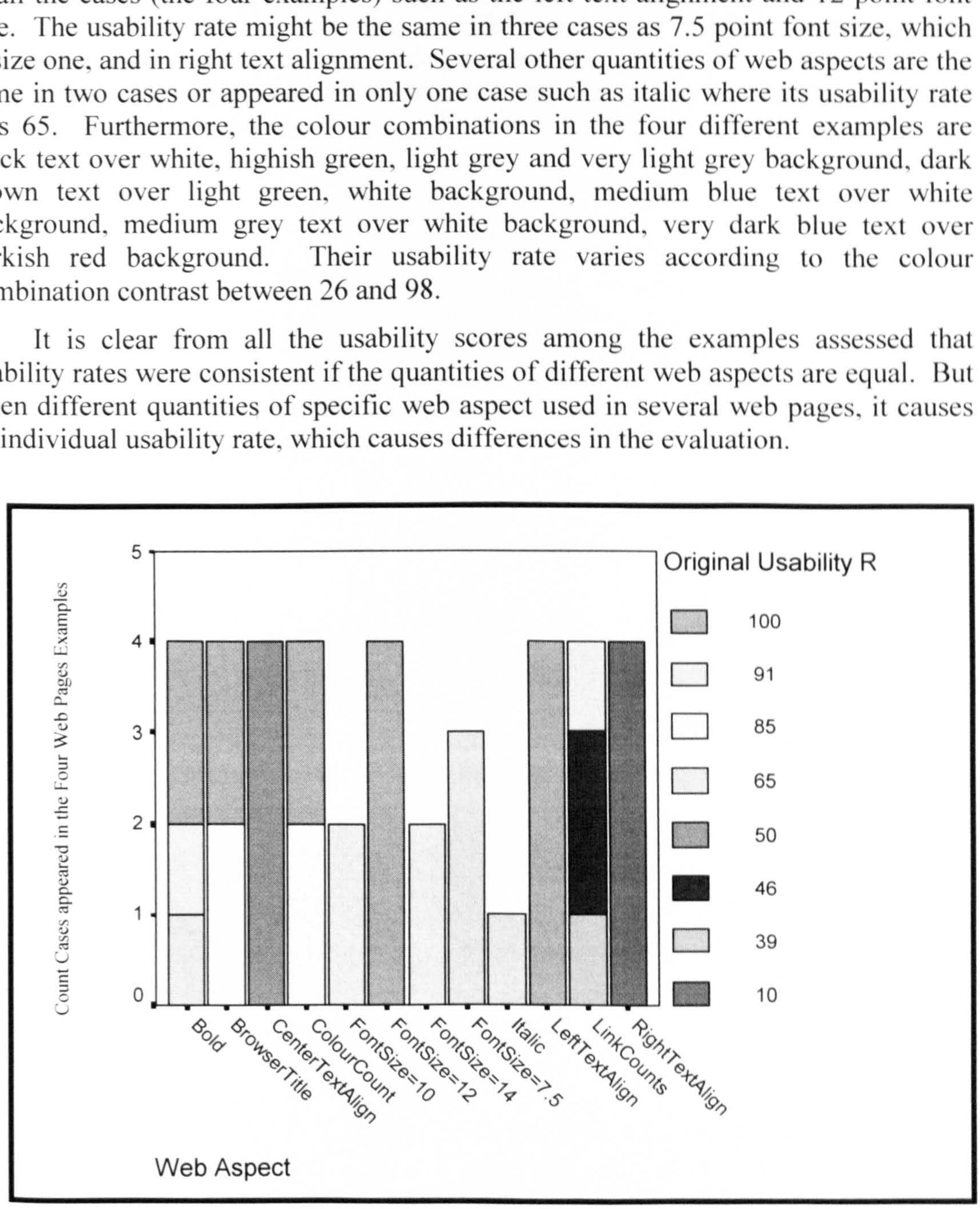


Figure 8.5: The Evaluation Analysis for Different English Web Pages Examples

8.10 Conclusion

This chapter demonstrated the web page analysis and evaluation performance performed by the analysis and evaluation phases respectively. The results reported from the last phase showed that the fuzzy model is sufficient to produced usability rates for the different web aspects included in the study. These usability rates are consistent with the current usability guidelines together with the user satisfaction (see Chapter 7). The findings produced by the fuzzy model signify a major step toward achieving the original goal of this research.

The examples' assessments illustrate the agent evaluation process and provide more insight into the dynamic web page aspects. Also, the examples prove that the characteristics of Arabic web pages differ from their English counterparts in predictable ways, which can be considered as another achievement of the study. The description of the three learning approaches and the differences between them are discussed in the next chapter.

Chapter Nine

Evaluating the Agent's Learning

9.1 Introduction

It is important for the web usability evaluation agent to have a certain degree of intelligence which allows the agent to customise its web usability assessment. The intelligence in learning the webmaster's preferences is the second type of customisation investigated in this study; the other type of intelligence was to customise the assessment rules to the users' tastes. Chapter 6, Section 6.5.3 describes the three learning approaches methodology utilised by this research to establish the agent's intelligence. The three learning approaches depend on the webmaster's feedback, but they differ in their implementation.

Chapter 8, Section 8.5 discussed the analysis and evaluation processes performed by the web usability evaluation agent. Aljazeera and BBC web pages were used to explain the agent's analysis and evaluation as an example of Arabic and English web pages respectively. The same web pages were used to test the three learning approaches. However, this chapter describes the customisation of the evaluation on the basis of webmaster's feedback and how these learning approaches influence the next evaluation.

9.2 Web Page Examples and the Webmaster's Feedback

As mentioned before, Aljazeera and the BBC are the examples to investigate the reaction of the three learning approaches. The three learning approaches were evaluated after the analysis and evaluation processes. The agent was run through 20 iterations to test its learning performance. In each iteration the web usability evaluation agent received different feedback from the webmaster for each web aspect. The same feedback was applied for all three approaches each time. However, each learning approach processes the webmaster's feedback differently. Table 9.1 and Table 9.2 show the different values of the webmaster's feedback for the web aspects tested as well as all the iterations in evaluating both Aljazeera and the BBC web pages respectively.

9.3 Agent's Learning Process

As mentioned in Chapter 6 (see Section 6.5.3) the learning process started after the first evaluation was produced by the agent and sent to the webmaster. In return the agent received the webmaster's feedback and used it to learn the webmaster's tastes. The following sections discuss the findings established by the three learning approaches for both Aljazeera and BBC web pages.

Web Metric Tested	Aljazeera Webmaster's Feedback																		
	Iteration Number																		
	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10	No.11	No.12	No.13	No.14	No.15	No.16	No.17	No.18	No.19	No.20
Font size = 12pt	31	30	12	30	16	51	38	55	47	38	88	88	80	92	98	88	97	85	91	89
Image sizes	34	15	16	15	27	39	63	45	39	69	83	99	97	96	88	91	93	79	89	96
Link count	7	22	11	32	30	53	42	49	67	39	85	96	88	78	99	86	75	93	86	88
Browser title length	15	30	21	8	30	36	51	51	55	57	82	97	76	93	91	84	97	93	95	93
Right text alignment	9	13	7	7	27	61	61	43	39	49	89	94	86	75	81	77	99	97	96	79
Centre text alignment	18	8	25	31	25	61	38	58	43	67	77	81	75	98	97	89	97	91	94	96
Left text alignment	15	16	25	18	9	46	54	62	67	37	78	99	95	96	83	98	85	78	85	87

Table 9.1: The Aljazeera Webmaster's Feedback During the 20 Iterations

Web Metric Tested	BBC Webmaster's Feedback																		
	Iteration Number																		
	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10	No.11	No.12	No.13	No.14	No.15	No.16	No.17	No.18	No.19	No.20
Font size = 7.5pt	25	28	17	10	19	43	62	40	58	51	90	86	96	94	77	97	89	77	91	90
Font size = 10pt	8	6	34	20	19	64	41	67	65	66	82	91	98	86	92	90	86	88	80	83
Font size = 12pt	27	33	22	32	14	61	42	55	45	53	91	91	91	88	91	87	97	93	88	77
Bold text	35	33	33	33	12	43	63	43	57	63	76	91	94	85	83	83	99	87	94	76
Image sizes	15	22	23	23	7	43	43	46	63	57	95	85	81	76	83	91	87	89	89	76
Link count	18	24	15	6	27	54	48	51	69	39	85	96	86	86	84	83	93	76	93	94
Browser title length	32	24	21	17	30	55	37	40	42	37	95	91	98	88	96	87	87	80	77	80
Right text alignment	19	24	33	28	13	43	57	41	56	55	94	98	77	77	95	89	89	78	81	80
Centre text alignment	33	30	19	20	13	55	63	62	51	45	89	98	78	98	98	83	93	94	76	75
Left text alignment	26	28	30	27	11	42	50	36	39	49	89	82	87	93	96	76	79	89	87	88

Table 9.2: The BBC Webmaster's Feedback During the 20 Iterations

Web Aspect	Webmaster's Feedback	Original Evaluation	Learning Approaches Evaluation		
			Fuzzy Average Learning	Fuzzy Average Learning	Q-learning Evaluation
Font size = 12	31	100	47	40	31
Links count	7	33	3	9	7
Browser title word count	15	98	20	13	15
Image sizes	34	0	35	42	34
Right text alignment	9	100	16	10	9
Centre text alignment	18	50	12	12	18
Left text alignment	15	10	6	13	15

Table 9.3: The Aljazeera First Iteration Three Learning Approaches Evaluation

9.3.1 Applying the Fuzzy Average Learning Approach

The foundation for determining the new usability rates for different web aspects tested in the fuzzy average learning approach is rooted in calculating both the average of the webmaster's feedback and the previous usability rate (see Chapter 6 Section 6.5.3.1). Both the webmaster's feedback and the previous usability rates were collected from profiles maintained by the controller phase (see Figure 6.1). These two profiles are updated when the agent is given a web page to test and consequently it receives feedback from the webmaster.

For the first iteration, the determination of the new usability rates produced by this approach is driven from earlier webmaster's feedback and the agent's original evaluation. The results of the fuzzy average approach first iteration is shown in the fourth column in Tables 9.3 and 9.4 for the Aljazeera and BBC respectively. It is clear that the new evaluation rates produced by this approach were close to the webmaster's feedback. The amount of change is based on the value of the webmaster's feedback in combination with the agent's original usability rate evaluation.

9.3.2 Applying the Fuzzy Learning Approach

The fuzzy learning approach, unlike the fuzzy average learning approach, is based on the last webmaster's feedback in determining the new usability rate. If the webmaster's feedback is minimised then the usability rate will be reduced and vice versa, as explained in Chapter 6 (Section 6.5.3.2). The last feedback can be taken from the webmaster's profile developed by the control phase of the agent.

It is clear that almost all the new usability rate findings produced by the fuzzy learning approach in the first iteration decreased, based on the webmaster's feedback. However, two usability rates were increased in both the Aljazeera and the BBC (see Table 9.3 and Table 9.4). It increased for the image sizes in both web pages, whereas the re-evaluation of the right text alignment in the BBC and the left text alignment in

Aljazeera since the webmaster's feedback in both of them is greater than the agent's original usability rates. In order to customise the evaluation to the feedback this fuzzy learning approach tuned the usability rate to the webmaster's feedback.

Web Aspect	Webmaster's Feedback	Original Evaluation	Learning Approaches Evaluation		
			Fuzzy Average Learning	Fuzzy Learning	Q-learning
Font size = 12	27	100	30	22	27
Font size = 7.5	25	65	25	22	28
Font size = 10	8	91	13	9	8
Bold text effect	35	91	46	35	42
Links count	18	100	18	12	18
Browser title word count	32	59	34	41	33
Left text alignment	26	100	30	22	26
Image sizes	15	0	6	13	14
Right text alignment	19	10	5	11	19
Centre text alignment	33	50	34	41	33

Table 9.4: The BBC First Iteration Three Learning Approaches Evaluation

9.3.3 Applying the Q-Learning Approach

The Q-learning approach can customise the calculation of the new usability rate to the webmaster's preferences by processing both the web metrics profile and the webmaster's feedback (see Chapter 6 Section 6.5.3.3). Each web metric identifies certain states that needs to be evaluated according previous feedback from the webmaster. The findings in the first iteration were the same as the webmaster's feedback, since all the Q initial values were zeros.

By comparing the usability rates calculated via this learning approach with the agent's original usability rate, depicted in Tables 9.3 and 9.4, there are huge differences among them. These differences were increased or reduced the original usability rates calculated by the agent and are influenced by the webmaster's feedback. The findings from the revaluation, of both the Aljazeera and the BBC web pages, showed a drastic change from the agent's original usability rates since this learning approach tries to modify the new evaluation to customise it to the webmaster's feedback.

9.4 Comparison Between the Three Learning Approaches

As can be seen from all the new usability rates computed in the first iteration with the three different learning approaches they all changed from the original usability rate. The changes appeared for every web aspects analysed for both the BBC and Aljazeera web pages illustrated in the Tables 9.3 and 9.4. All these changes are affected by the webmaster's feedback but as can be seen from the first iteration in

Tables 9.1 and 9.2, the feedback range between (7- 35) disagrees and strongly disagrees, so most of the new usability rates are reduced.

However, more than one iteration is needed to test the effectiveness in the customisation over the three learning approaches. As mentioned in Chapter 6 twenty iterations were performed to study the three learning approaches. The analysis of the responses from all the learning approaches is discussed in the next section.

9.4.1Responses

From the first iteration analysis it is obvious that the usability rate computed by the three learning approaches was close to the first webmaster feedback. However, in order to investigate the effectiveness of each learning approach for the long run, it is necessary to examine their performance over several iterations and compare their output with the feedback. By running the agent for twenty iterations for each web page Aljazeera and BBC different reactions were collected. Appendix *B* contains tables with details of the three learning approaches responses for the Aljazeera and the BBC web pages tested.

9.4.1.1 Aljazeera Learning Approaches Responses

The webmaster's feedback values varied during the running of the agent over all the possible responses. The usability evaluation was customised to the webmaster's feedback via the three learning approaches. Because the web page evaluation is based on individually assessing different web aspects, the new scores for the usability evaluation is independently customised.

Figure 9.1 shows three learning approaches responses to the font size over the 20 iterations. The left hand side of Figure 9.1 shows that for the first few iterations the webmaster's feedback was low and then it starts to increase representation in the right hand side of the figure. It is clear that these changes in the feedback were associated with improvement in the evaluation produced by all of the three learning approaches. Table 9.5 depicts the 12 point font size evaluation associated with the webmaster's feedback for the 20 iterations.

The effects of the webmaster's feedback were similar for all web aspects analysed and evaluated by the usability evaluation agent. Appendix *B* contains the figures that illustrate the changes of the evaluation based on the webmaster's feedback improvement for the other web aspects analysed from the Aljazeera web page. The next section explains the changes in the BBC web page evaluation.

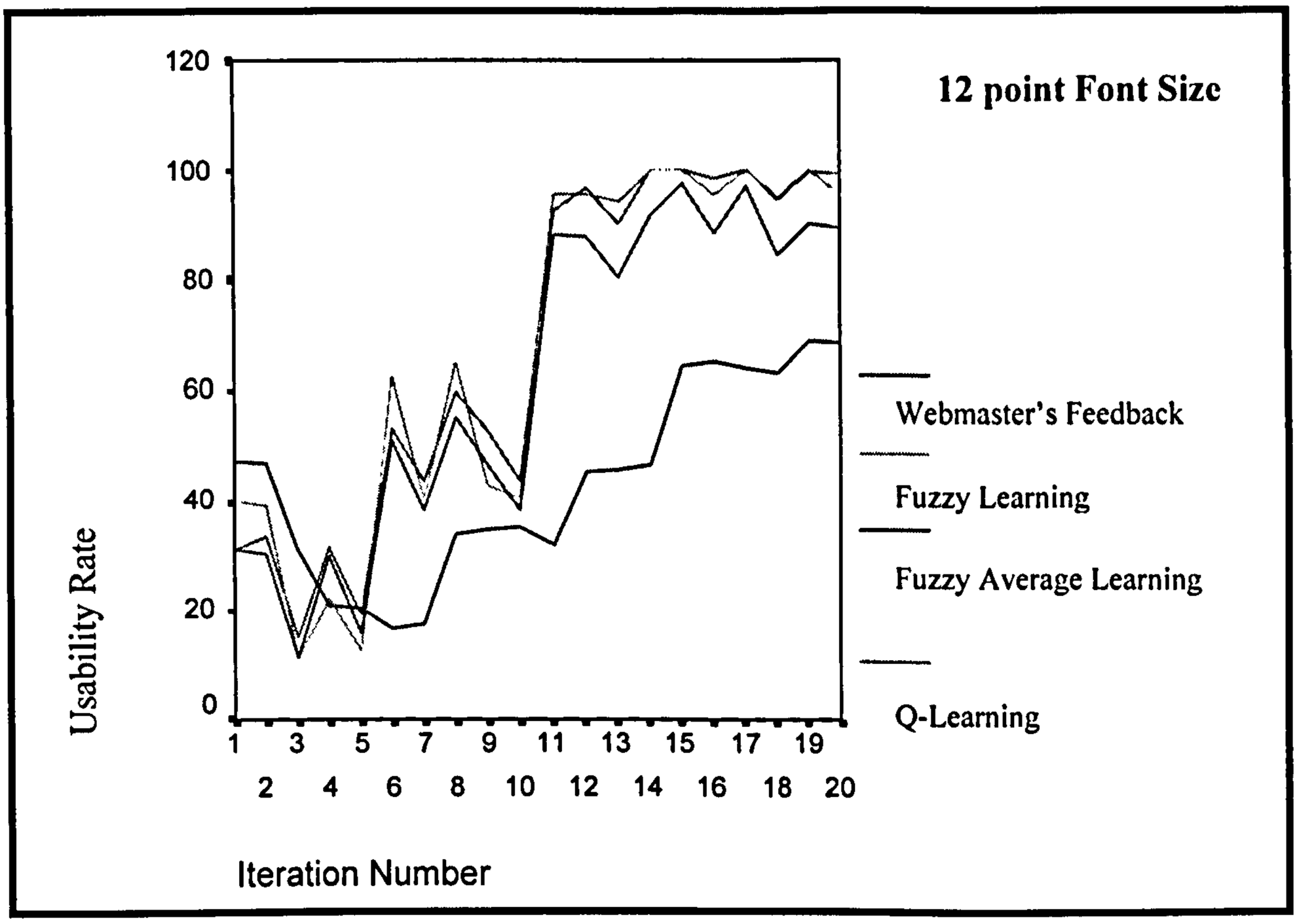


Figure 9.1: The Customisation of the Font Size (12 Point) UsabilityRate in Aljazeera Web Page

Iteration Number	Webmaster's Feedback	Usability Evaluation Learning Approaches		
		Fuzzy Average	Fuzzy Learning	Q-learning
1	31	47	40	31
2	30	47	39	33
3	12	31	11	15
4	30	21	22.1	31
5	16	20	13	19
6	51	17	62	53
7	38	18	40	44
8	55	34	65	60
9	47	35	43	53
10	38	35	40	44
11	88	32	96	93
12	88	45	96	97
13	80	46	94	90
14	92	46	100	100
15	98	64	100	100
16	88	65	96	98
17	97	64	100	100
18	85	63	95	95
19	90	69	100	100
20	89	68	96	100

Table 9.5: The Changes in Aljazeera Font Size (12 Point) Evaluation

9.4.1.2 BBC Learning Approaches Responses

As mentioned in Chapter 6, Section 6.5.5 the webmaster's feedback values were varied during the running of the agent. Based on the webmaster's replies the new usability evaluation would be changed. Customising the usability evaluation to the webmaster's feedback was performed by the three learning approaches. In each iteration the same webmaster's feedback was used for the three learning approaches.

Three different font sizes were analysed from the BBC web page (see Chapter 8 Table 8.3), and then three individual evaluations were produced using three independent webmaster's feedback for each size. Figure 9.2 shows the three learning approaches responses for 7.5 font size using the changes in the webmaster's feedback over the 20 iterations, while Figure 9.3 and Figure 9.4 are the responses for the 10 point and 12 point font sizes respectively. Each figure is associated with table that represents the different values of the webmaster's feedback and the three usability rates after the three learning techniques, fuzzy average, fuzzy and Q-learning.

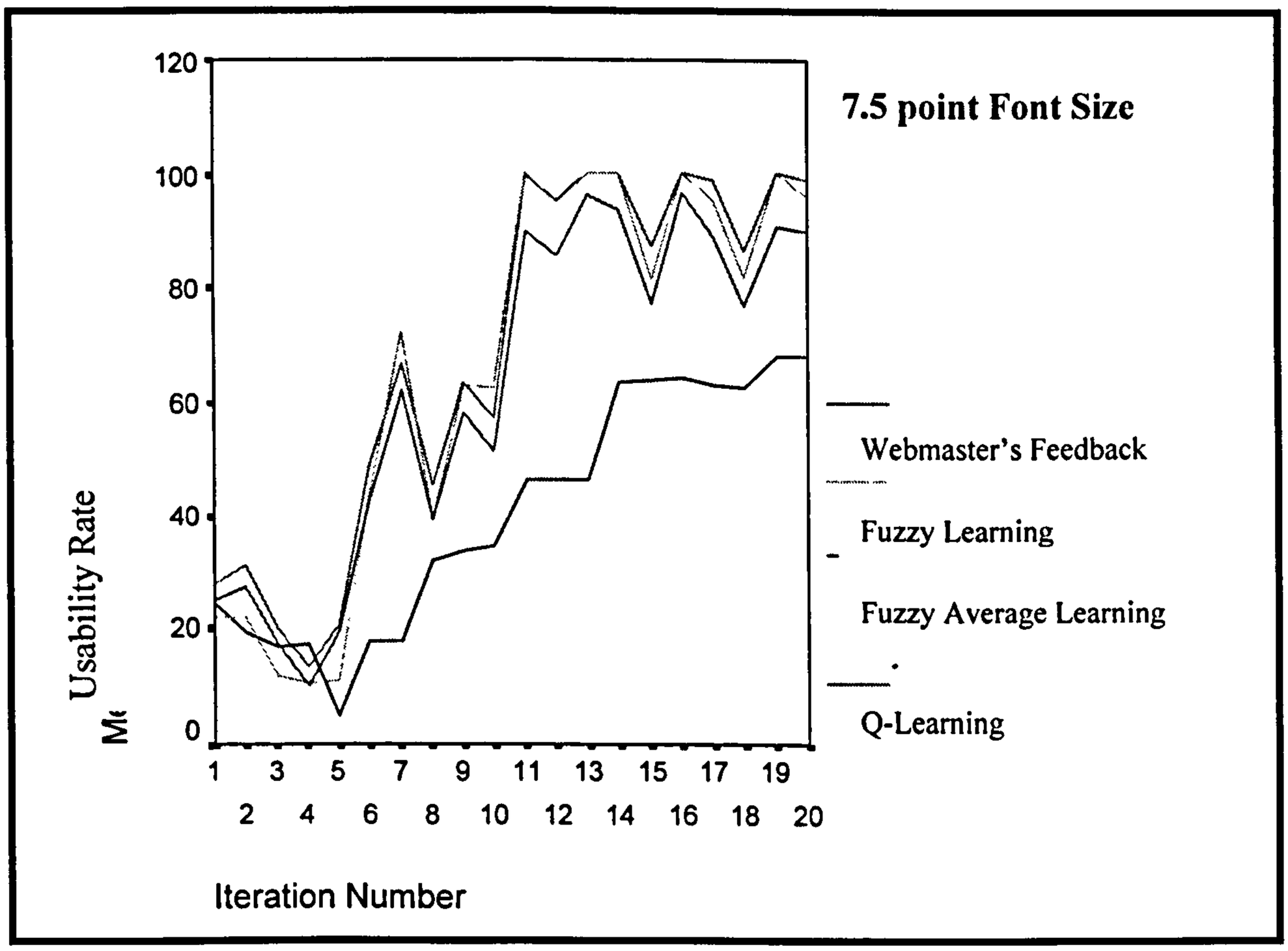


Figure 9.2: The Customisation of the Font Size (7.5 Point) Usability Rate in BBC Web Page

Iteration Number	Webmaster's Feedback	Usability Evaluation Learning Approaches		
		Fuzzy Average	Fuzzy Learning	Q-learning
1	25	25	22	28
2	28	20	22	31
3	17	17	12	20
4	10	17	11	14
5	19	5	11	21
6	43	18	44	49
7	62	18	72	67
8	40	32	40	46
9	58	34	63	63
10	51	35	63	57
11	90	46	100	100
12	86	46	95	96
13	96	46	100	100
14	94	63	100	100
15	77	64	82	87
16	97	64	100	100
17	89	63	96	99
18	77	62	82	87
19	91	68	100	100
20	90	68	96	99

Table 9.6: The Changes in the BBC Font Size (7.5 Point) Evaluation

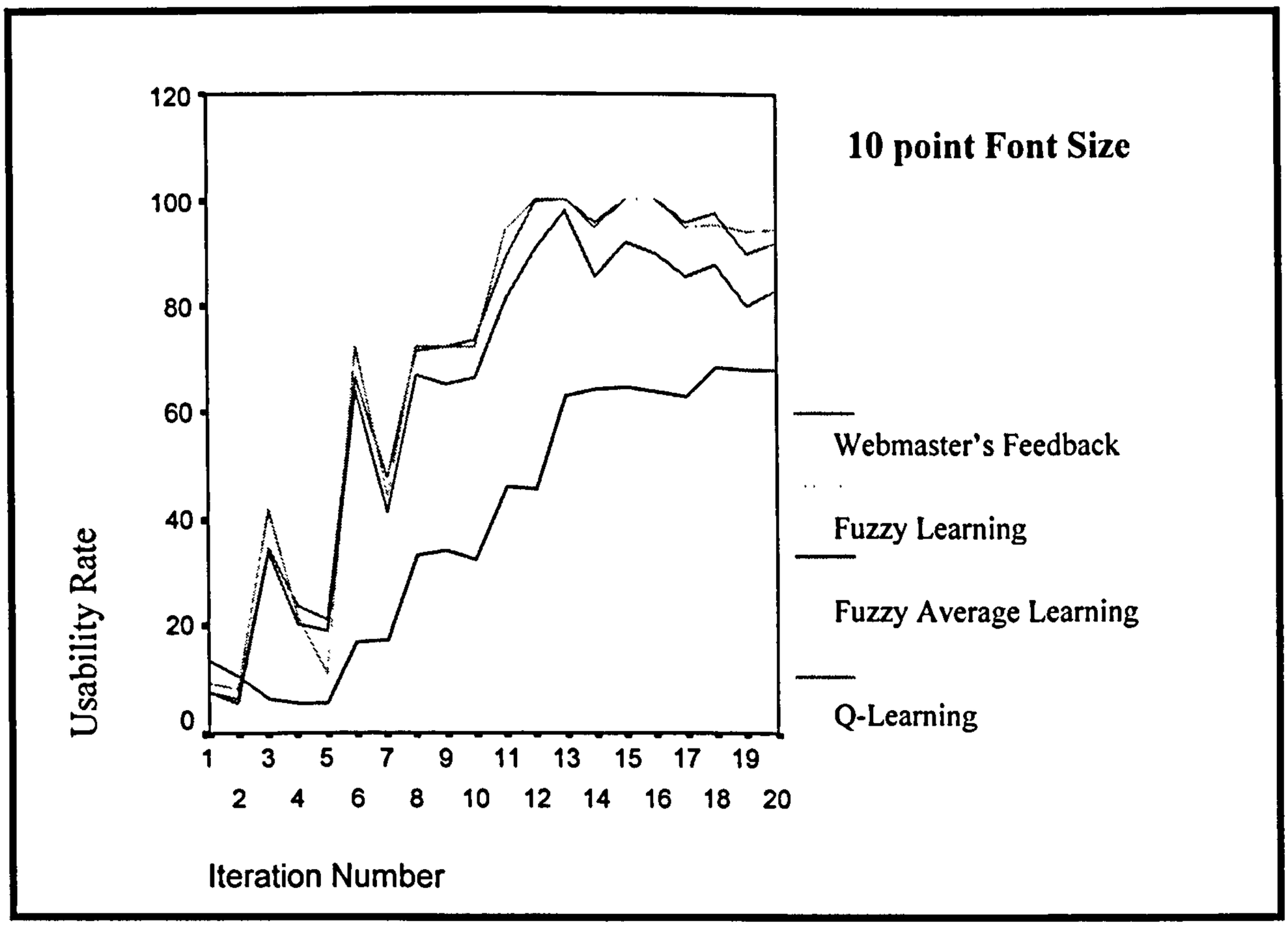


Figure 9.3: The Customisation of the Font Size (10 Point) Usability Rate in BBC Web Page

Iteration Number	Webmaster's Feedback	Usability Evaluation Learning Approaches		
		Fuzzy Average	Fuzzy Learning	Q-learning
1	8	14	9	8
2	6	11	8	6
3	34	6	42	35
4	20	6	22	24
5	19	6	11	21
6	64	17	72	66
7	41	17	44	48
8	67	33	72	72
9	65	34	72	72
10	66	32	72	74
11	82	46	95	89
12	91	46	100	100
13	98	63	100	100
14	86	64	95	96
15	92	65	100	100
16	90	64	100	100
17	86	63	95	96
18	88	69	96	98
19	80	68	94	90
20	83	68	95	92

Table 9.7: The Changes in the BBC Font Size (10 Point) Evaluation

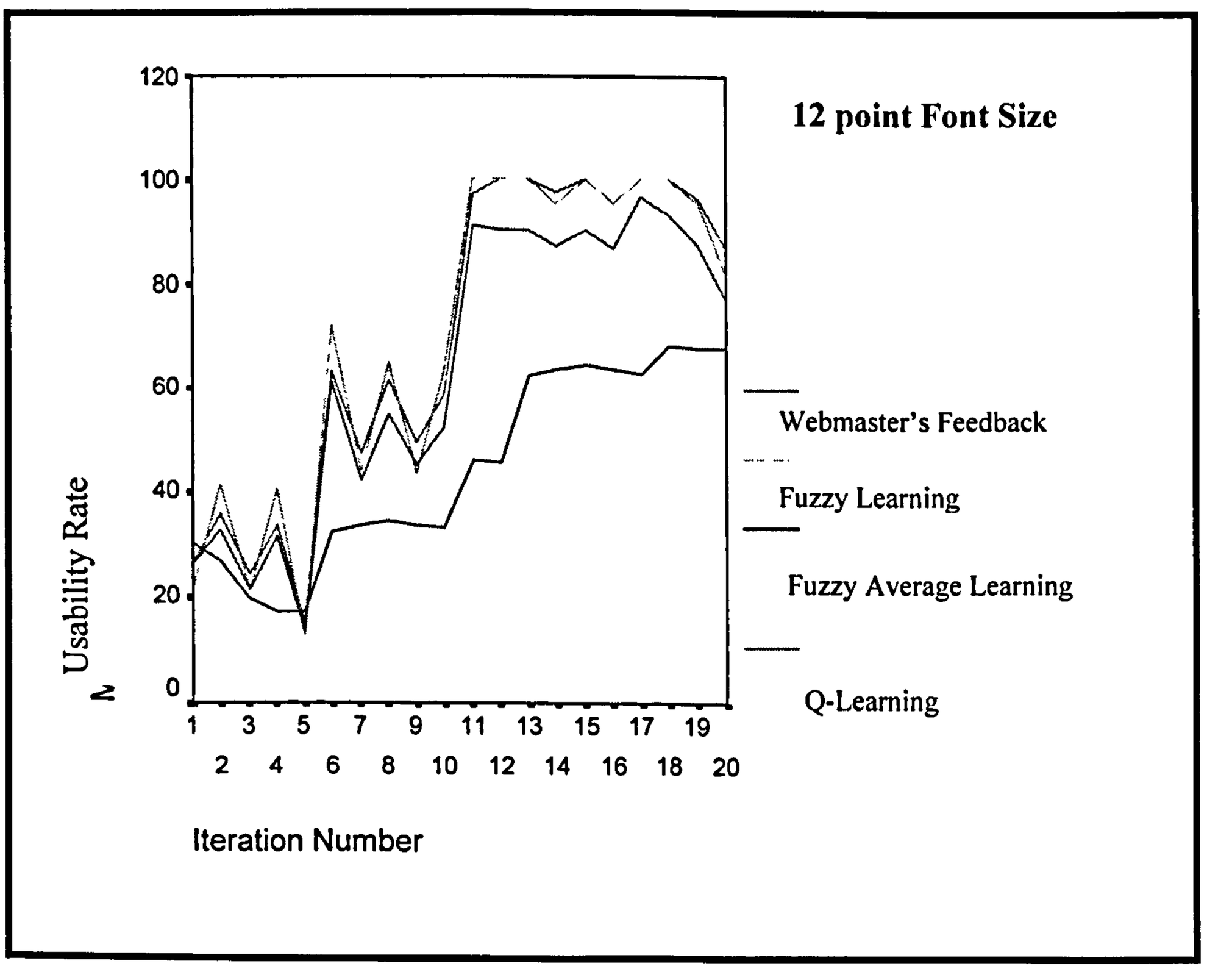


Figure 9.4: The Customisation of the Font Size (12 Point) Usability Rate in BBC Web Page

Iteration Number	Webmaster's Feedback	Usability Evaluation Learning Approaches		
		Fuzzy Average	Fuzzy Learning	Q-learning
1	27	30	22	27
2	33	27	41	36
3	22	20	21	25
4	32	17	40	34
5	14	17	13	15
6	61	33	72	63
7	42	34	44	47
8	55	35	65	62
9	45	34	43	50
10	53	33	63	59
11	91	46	100	97
12	91	46	100	100
13	91	63	100	100
14	88	64	95	98
15	91	65	100	100
16	87	64	95	96
17	97	63	100	100
18	93	69	100	100
19	88	68	95	96
20	77	68	82	87

Table 9.8: The Changes in the BBC Font Size (12 Point) Evaluation

The left hand side of all the BBC font sizes figures (Figures 9.2, 9.3, 9.4) indicates that the first few iterations of the webmaster's feedback were low and then they start to increase as shown in the right hand side of the figures. In each iteration, the usability scores produced by different learning approaches were affected by the webmaster's feedback. Table 9.6, Table 9.7 and Table 9.8 depict the 7.5, 10, and 12 point font size evaluation associated with the webmaster's feedback for the 20 iterations.

Each web aspect tested had individual feedback which reflected the webmaster's acceptance of the evaluation. Appendix *B* contains the detail of other web aspects analysed from the BBC web page and evaluated by the agent which has figures to illustrate the reflection of the webmaster's feedback on the usability scores calculated in each iteration.

9.5 Determining the Learning Approaches Accuracy

From all the figures in last two sections (Aljazeera and BBC learning approaches responses), it is clear that the evaluation produced from all the learning approaches reduced as the feedback reduced and grew as the webmaster's responses increased. However, not all the three learning approaches responses were the same, since it is clear that the fuzzy average learning approach evaluations were the farthest away from the webmaster's responses. Because the fuzzy average learning approach is based on the average webmaster's feedback which seems to be close to the webmaster's feedback during the first few iteration it is not effective for long runs of usability evaluation especially if drastic responses appear.

On the other hand, using the other techniques the evaluation produced by the fuzzy learning approach and the Q-learning approach were both consistent with the

webmaster's feedback. A drop in the webmaster's feedback values is associated with a reduction in both the fuzzy learning and Q-learning evaluations.

9.5.1 Learning Approaches Results Accuracy

It is realised from the figures illustrated earlier and the figures shown in Appendix *B* that the Q-learning and the fuzzy learning produce the closest usability evaluation to the webmaster's feedback. But the Q-learning responses appeared to be more identical to the webmaster's feedback changes. Also, it is obvious from all the figures that the Q-learning curve is consistent with the webmaster's feedback curve. In other words, it follows the details changes accurately and never goes below the webmaster's feedback. The Q-learning evaluation is either compatible or above the webmaster's feedback. However, the fuzzy learning evaluation does not show the same consistency as the Q-learning, because in some iterations the evaluation is below the feedback and in other iterations the evaluation appeared to be grater. Thus, the Q-learning evaluation exhibited more accuracy and consistency than the other learning approaches.

A standard multiple regression analysis was used in order to detect the most accurate learning technique that produced the closest usability evaluation to the feedback. The regression measurement is based on selecting the webmaster's feedback as the dependent variable, while the three learning approaches were independent variables.

It is clear that there were differences in the usability scores among the three learning approaches, but the best of all is the Q-learning approach. Q-learning gives the most reliable customised evaluation, based on the R square change. Table 9.9 illustrates the R square changes among all the approaches. However, fuzzy learning comes in the second significant learning approach while the least accurate learning approach was the fuzzy average learning.

In spite of the similarities in the usability evaluation for both the Q-learning and fuzzy learning approaches, the regression test shows that there is a huge difference in the R-square between them. This is attributed to the fact that the difference between the fuzzy learning's usability evaluation and the webmaster's feedback can be a positive number, a negative number or a zero, which significantly affected the regression test results. However, the difference between the Q-learning's usability evaluation and the webmaster's feedback appeared to be always negative. The Q-learning approach appeared to be greater than the webmaster's feedback, because it always maximise the reward of the webmaster.

Learning Approach	R Square	R Square Change	F Change
Q-Learning	0.993	0.993	63127.944
Fuzzy Learning	0.994	0.001	38.708
Fuzzy Average Learning	0.994	0.000	14.454

Table 9.9: Evaluating the Three Learning Approaches by Using the Simple Regression

9.5.2Implementation

As mentioned in Chapter 6 (Section 6.5.3), the implementation of each learning approach differs from the other approaches. All of them require knowledge of the web aspect state (current web metric measure) that needs to be evaluated, but they differ in the other data needed to determine the new evaluation. For instance, the fuzzy average approach is based on all the previous usability rates for the web aspect tested in addition to a collection of all previous webmaster's feedback for the web aspect evaluated. The fuzzy learning approach is founded only on the webmaster's last feedback, while the Q-learning approach is driven from the identification of the last feedback from the webmaster, together with all the previous evaluations of the same state of the web aspect.

It is clear that fuzzy learning approach needs less storage since it relies on the last webmaster's feedback only. The Q-learning approach needs the profile of usability rate produced for different web aspects while the fuzzy average approach needs more profiles to be maintained, all previous webmaster's feedback and all the other usability rates produced by this approach. Also, the fuzzy average learning approach involves 45 fuzzy rules to infer the effects that cause the changes in the fuzzy evaluation model and then the overall usability evaluation produced. But the fuzzy learning approach inference section influences the fuzzy evaluation model with only five fuzzy rules. The deduction of the new usability rate by using the Q-learning approach is based on the formula presented in Chapter 6 (Section 6.5.3.3).

9.6 Conclusion

This chapter attempted to answer questions such as: is it possible to customise the usability evaluation, how effective are customised usability rates produced and which is the most efficient learning approach? Thus, this chapter presented the results from the three learning approaches (fuzzy learning, fuzzy average learning and Q-learning) tested for 20 iterations by using the Aljazeera and the BBC web pages. The detailed description of the first iteration for all the learning approaches was used to give an insight into the reactions of each approach individually.

The study demonstrated that it is possible to customise the usability evaluation through all three learning approaches. Each of the learning approaches provided different customisation effects.

However, the consistency of the evaluation results customised with the webmaster's feedback was measured by applying a standard regression test. The results showed that the Q-learning approach was significantly better than the other learning approaches. It is the closest approach to the webmaster's feedback, followed by the fuzzy learning approach while the weakest customised evaluation was produced by the fuzzy average approach.

Chapter Ten

Conclusion and Future Work

Users from all over the globe can be attracted by WWW services. Both WWW vendors and developers invest large amount of time in increasing the web's productivity. During the last decade, abundant usability guidelines and recommendations have been created, each of them aimed at improving web design. Chapter 2 presented an intensive survey of usability guidelines under scrutiny together with other recommendations which contribute to develop legible web pages. Furthermore, several automated and non-automated usability methods have been proposed and are discussed in Chapter 3. These methods

- Detect problems in the design that might prevent the users from achieving their goals.
- Improve the design by finding more effective and efficient alternatives to the existing one.
- Discover new user interests or preferences which might cause changes in the usability guidelines.
- Become aware of the users' behaviour and manner of performing certain tasks in order to consider them in future designs.

Automated usability techniques are considered as an essential support to non-automated techniques and have several advantages over them as described in Chapter 3. Empirical automated usability evaluation techniques find a design's mistakes in a shorter time and at much less costs than non-automated. Examples include WebTango (Ivory, 2001), WebRemUsine (Paganelli and Paterno, 2002; Paterno, 2003) and others discussed in Chapter 3. But these automated techniques, despite their variety and abilities, are not able to customise their evaluation rules to match unpredictable changes in the usability guidelines or user requirements, and they cannot assess web pages implemented in languages other than English. In other words, the main weaknesses in the previous automated techniques are lack of customisation and intelligence.

The agent technique is a recently developed method that is capable of adjusting its behaviour to its owner's preferences and user's needs. This technique was described in Chapter 4. Thus, it is appropriate to utilise this technique to develop an adaptable automated usability evaluation tool. In order for the agent to be adaptable it should learn either from the environment or from its owner or users. Chapter 4 also includes the descriptions of several learning methods that are used to provide the agent with this ability.

This research is based on two usability evaluation methods. The first is user satisfaction (see Chapter 5) and the second is to heuristically adapt a group of existing guidelines to control the web page evaluation (see Chapter 6). The user satisfaction was performed to identify the Arabic users' preferences and is important in determining the basics of usability guidelines suitable for assessing Arabic web pages. Additionally, it was used to validate the usability guidelines for English web pages.

This was necessary because users' tastes change over the time, and because there were conflicts in some of the existing guidelines, and also to limit some of imprecise guideline measures. The findings from the user satisfaction in combination with the current usability guidelines are the foundation for the heuristics evaluation performed by the second part of the study.

In spite of this research achievement in identifying the key usability guidelines for Arabic and validating the current English usability principles, more investigations should be performed to discover additional usability guidelines keeping in mind several aspects such as users' demographics and users' psychographics. Both of them significantly affect the users' preferences as well having cultural effects. Additionally, other aspects should be investigated in more details. For instance, in order to identify the specific number of links on a web page, the question should specify the type of the web page. In other words, the participants should have an idea about the class of the web page; whether it is a home page, an index page or a text page. Furthermore, the other questions dealing with colours would be answered more accurately if the colours were illustrated with the questions.

The second section of this study utilised agent technology to perform an evaluation on both Arabic and English web pages. However, the agent's intended goals are not limited to heuristically evaluating the web pages but also to customising the evaluation to its webmaster's preferences. These two major goals have not been achieved by earlier automated usability evaluation techniques.

The multilingual automated web usability evaluation agent customisation is not restricted to the webmaster's needs, instead it can be customised to web usability experts recommendations. This substitution shows that the agent approach has a variety of utilisations, which can help in increasing the web productivity and assisting web experts and researchers. Nothing needs to be changed in the evaluation agent construction when substituting the usability specialist for the webmaster, even the initial agent evaluation is the same, since the basic assessment is driven by the current usability guidelines in combination with the user satisfaction findings. However, the agent assessment customisation could be varied as the usability specialist requires. The differences would appear in customising the evaluation to match of each of their needs, because they might have other guidelines which contradict the agent's usability rules.

The user satisfaction findings pointed to differences between two dissimilar cultures: English and Arabic. To reach a certain degree of stability and consistency in evaluating web pages with dissimilar languages, it is necessary to customise the evaluation.

There were notable cultural differences between English and Arabic users and the different scripts affect the readability and the users' preferences too, since the evidence showed that Arabic users prefer larger font sizes especially for the presentation of important text. But the English users prefer a smaller font size and prefer bold font to emphasise important text. Both Arabic and English users prefer a short document title length. There is a significant difference in the number of links within an attractive web page since Arabic users dislike many links in their web pages unlike their English counterparts. The text and heading alignments are consistent with both languages writing directions. Arabic users significantly prefer right alignment direction, while English users prefer left alignment.

All of these differences were applied within the fuzzy rules that control the agent's evaluation. The results reported from the fuzzy model produced efficient usability rates for the different web aspects included in the study since these usability rates are consistent with the current usability guidelines together with the results of user satisfaction. The sufficient findings produced by the fuzzy model signify a major step toward implementing an intelligent usability evaluation approach, which is one of the original goals of this research.

This study represents an important step towards automating the customisation of the web usability evaluation, which other automated usability evaluation tools cannot perform. All the learning approaches are consistent with the webmaster's needs, but the Q-learning approach is the most reliable, of the three learning approaches, in its customisation to the webmaster's preferences. Establishing an intelligent automated usability evaluation agent that can customise its evaluation to the webmaster's preferences was one of the achievements of this research.

At this stage the multilingual automated web usability evaluation agent can measure the readability criteria, of both Arabic and English web pages, as part of web usability. It can be utilised to serve webmasters as well as web usability specialists in finding out if the web pages tested is consistent with the current web usability guidelines together with the user satisfaction results. For instance, if Arabic or English usability specialists want to check several web aspects such as font size, text alignment, different text effects and colour contrast, then they can use the multilingual automated web usability evaluation agent.

The analysis of the web page introduced in this research is based on the web page over all or as one unit. For instance, there is no difference in assessing the small font size whether it appears on the top or the bottom of the web page. In another words, the usability evaluation rate is always low. However, in order to develop a more precise automated web usability evaluation tool that involves other criteria not included in this research such as assessing the font size based on its location on the web page, the web page analysis should be divided into several sections to increase the accuracy of the assessment. In this way the font size usability evaluation should be based on its position, for example the small font size usability evaluation rate should not be low if it appears on the bottom of the web page, while small font size should be given lower usability rate if it is displayed on the top. It is important to develop automated web usability evaluation tools that take into account the relationships between different web aspects such as the font size and its location on the web page and the utilisation of different colours and their positions.

The multilingual automated web usability evaluation agent does not have the capabilities to assess other usability criteria such as the relationship between different web aspects. For example, any the text effects such as bold, italic, underlined and the text colour, which influence the readability. Other usability guidelines responsible for evaluating the utilisation of more than one aspect, in other words, the guidelines that can measure the relationship between different web aspects, which might interfere with each other, should be considered. Thus, in order to develop an impartial and comprehensive usability evaluation tool that can assess every aspect within the web page, more usability guidelines should be included. The fuzzy rules for the current evaluation model should be extended to involve the additional web usability guidelines.

Even though this study established the basic usability guidelines for presenting legible Arabic web pages (on basis of the user satisfaction), more investigation should be performed to extend the Arabic usability guidelines. Expanding the Arabic usability guidelines should involve other web page and site aspects not included in this study, such as accessibility and navigation. Establishing usability guidelines is necessary for automated and non-automated web usability evaluation.

After a comprehensive review of the usability literature, the only guidelines existing are those related to the web pages written in English. Some of these guidelines can be applied to other languages, particularly those using Latin alphabets. In general there is a lack of usability guidelines for the presentation of different scripts such as Indian, Chinese, Korean or Japanese. There are no written guidelines for the majority of the world's scripts. Because of cultural differences, it is necessary to develop a set of usability guidelines for each of them. Different cultures guidelines need to be established to help in constructing successful multilingual web pages and to increase the web's productivity.

Another possible extension to this study is to integrate a multi-agent system to sense the changes in the web pages, evaluate and then improve them. This could be achieved by first developing an agent that senses the changes in the web pages under a certain server. Secondly, utilising the multilingual web usability evaluation agent established by this study to evaluate the web pages needs to be assessed. It is possible to enable the evaluation agent to evaluate the non-HTML pages. The third part is to develop an agent that can automatically update or improve the web page evaluated to the agent's usability rules, based on the webmaster's preferences.

Appendix A

User Satisfaction (Arabic and English) Findings' Comparison

Chapter 5 discusses the user satisfaction methodology, and the user satisfaction findings were discussed in Chapter 7. This section presents the full bar graphs which comparing the findings of Arabic speakers with their English counterparts.

A.1 Colour User Satisfaction Findings Comparisons

Colour user satisfaction is implemented to study the relationship between the colours usability and several concepts such as attractiveness and readability.

A.1.1 Attractiveness of the Four Selected Background Colours

The colour user satisfaction contains four online examples with different colour combinations and different colour backgrounds: light blue, black, mild light pink and medium green shades.

A.1.1.1 Blue Background Attractiveness

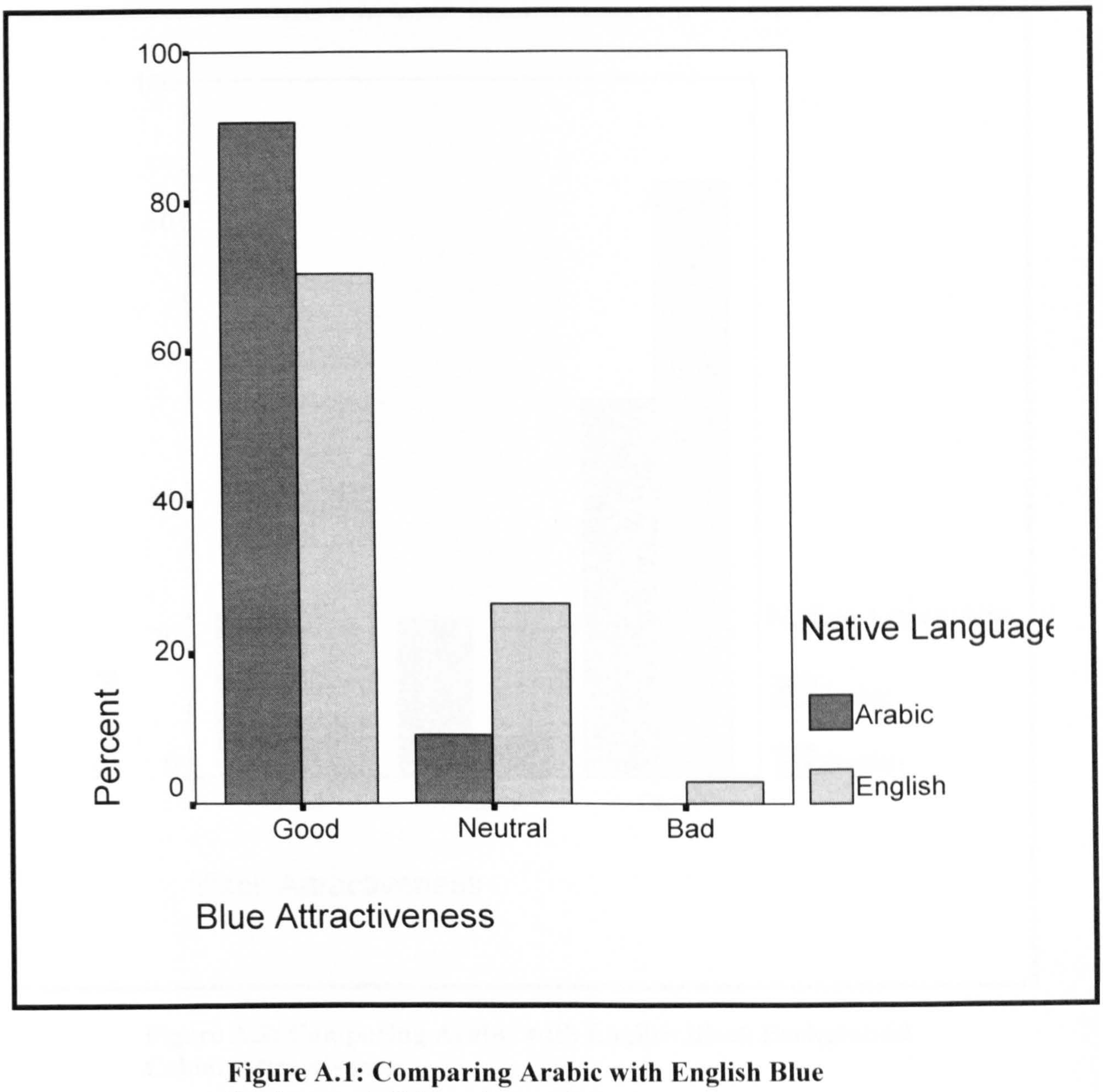


Figure A.1: Comparing Arabic with English Blue Background Colour Attractiveness

A.1.1.2 Black Background Attractiveness

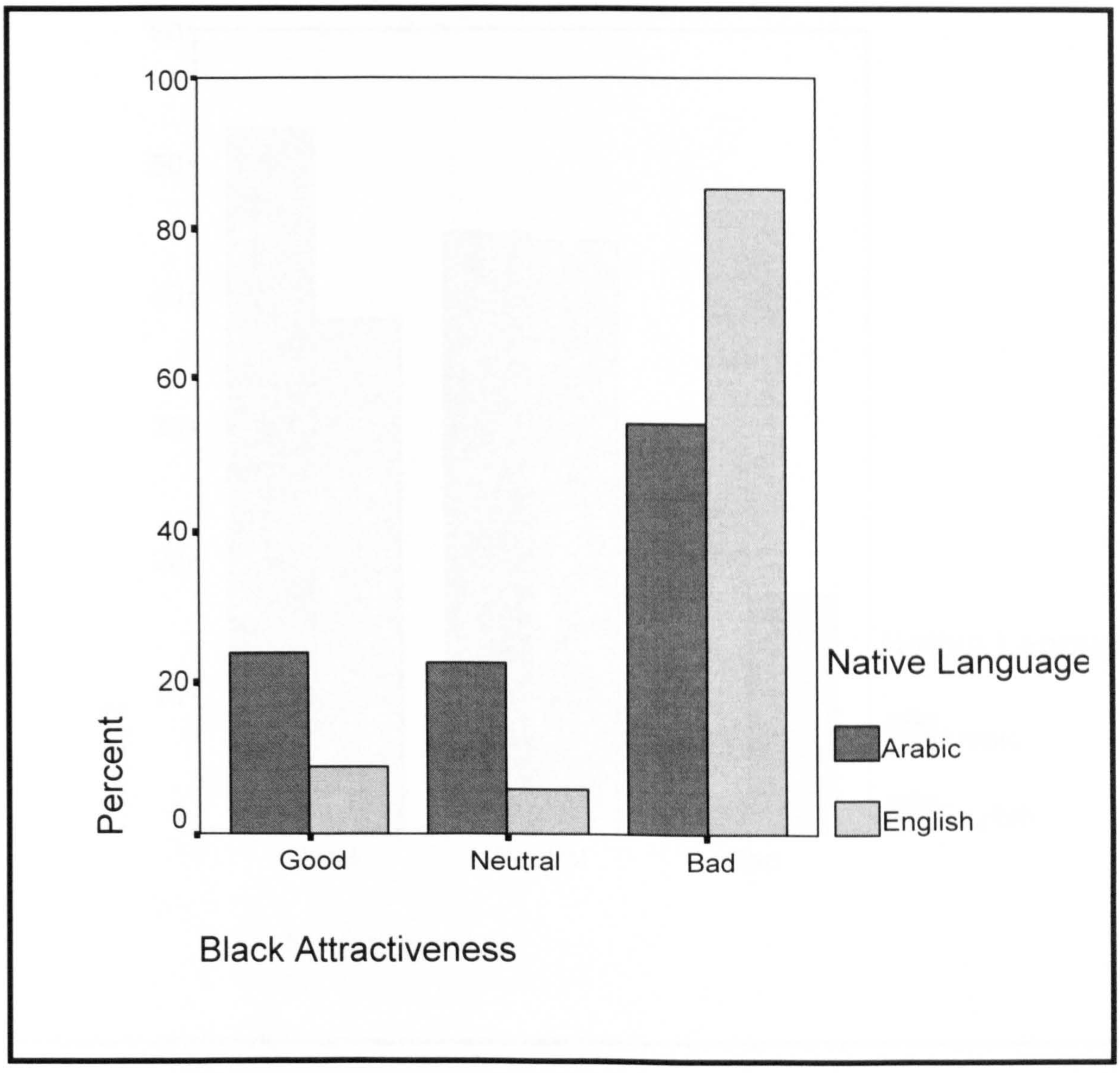


Figure A.2: Comparing Arabic with English Black Background Colour Attractiveness

A.1.1.3Pink Background Attractiveness

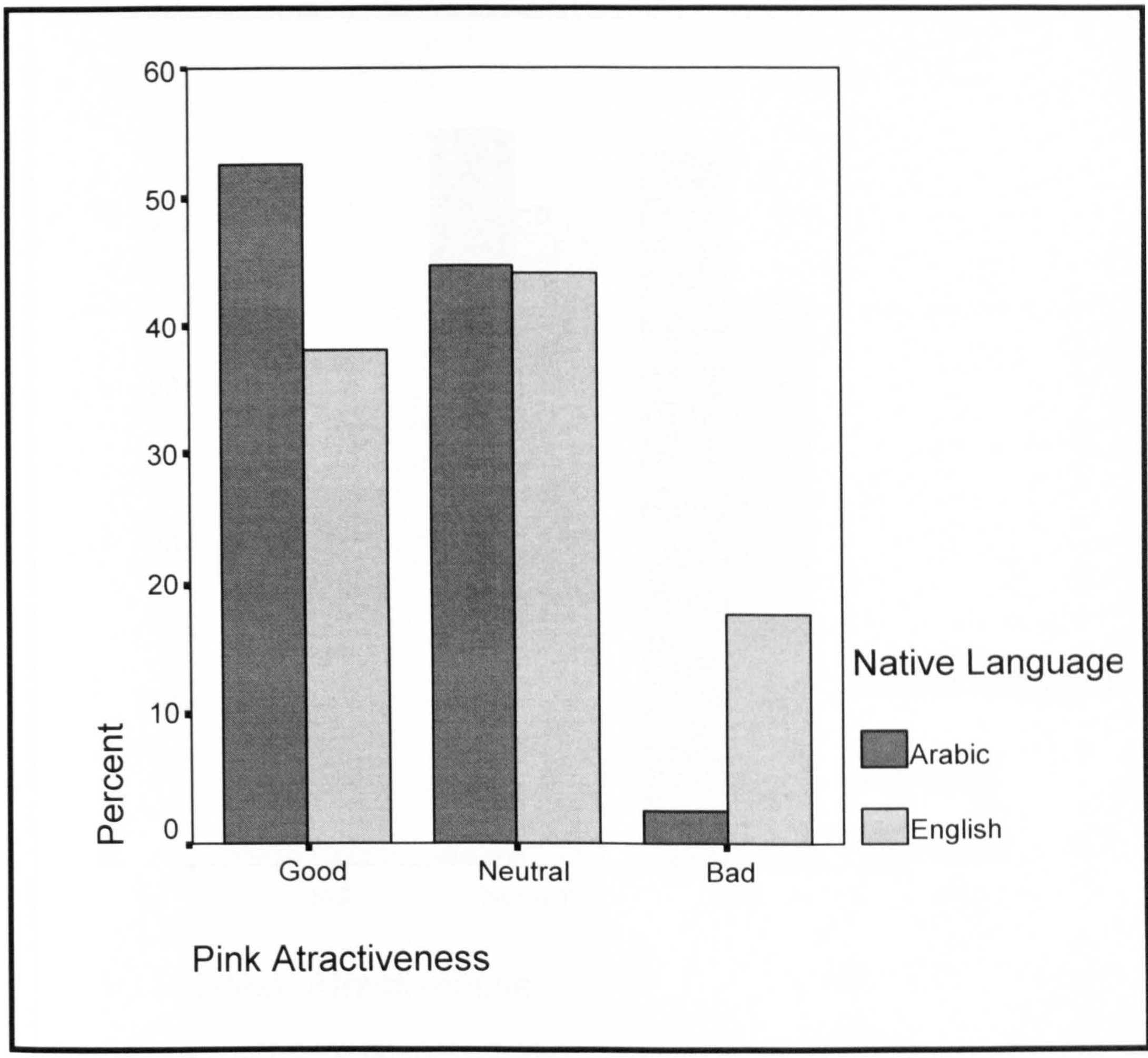


Figure A.3: Comparing Arabic with English Pink Background Colour Attractiveness

A.1.1.4 Green Background Attractiveness

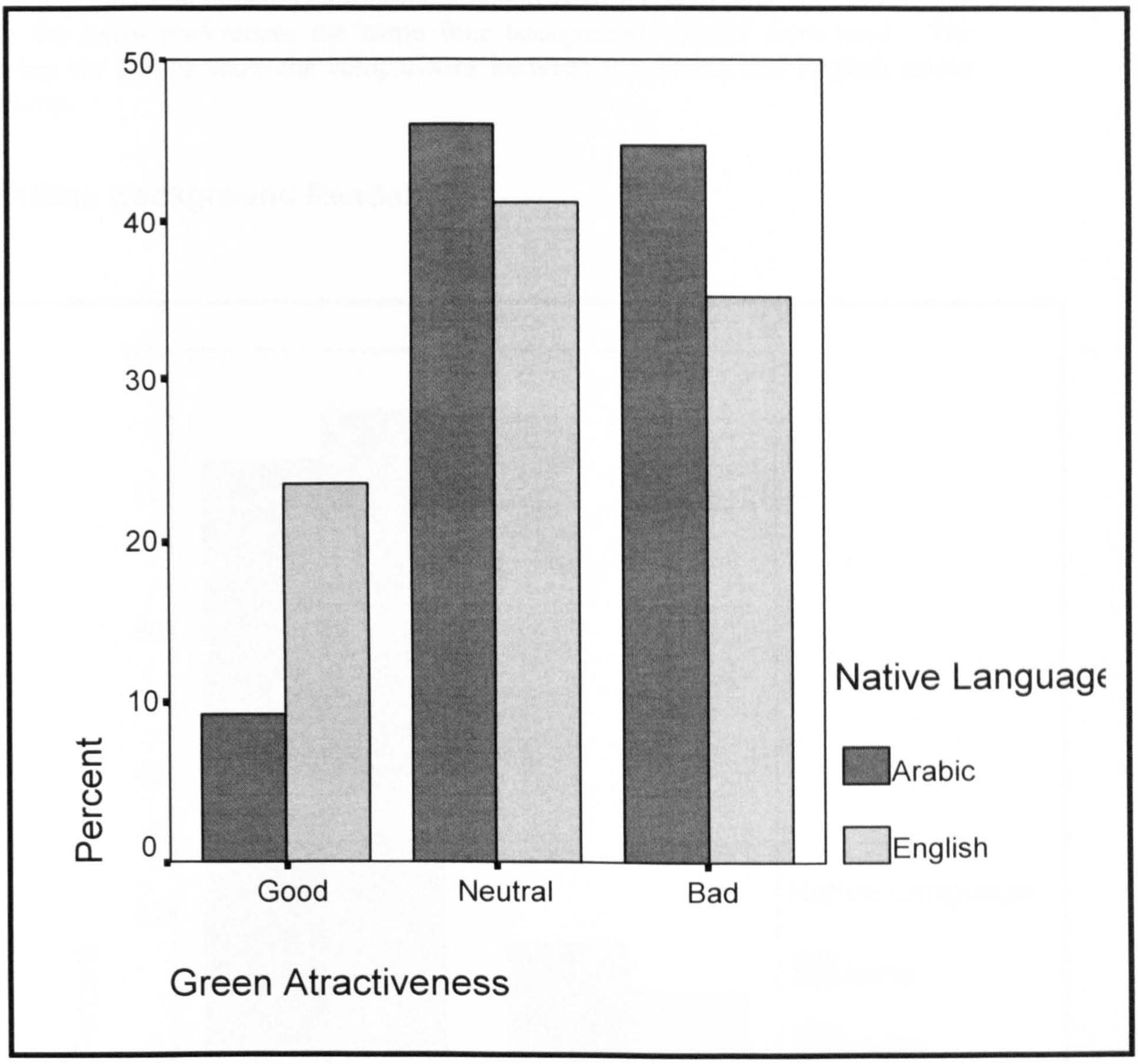


Figure A.4: Comparing Arabic with English Green Background Colour Attractiveness

A.1.2 ***Readability of the Four Selected Background Colours***

The colour combinations affect the readability of the web page, but in order to realise the users preferences the same four background colours were used. The following bar graphs show the comparisons between the Arabic and English colour readability.

A.1.2.1 Blue Background Readability

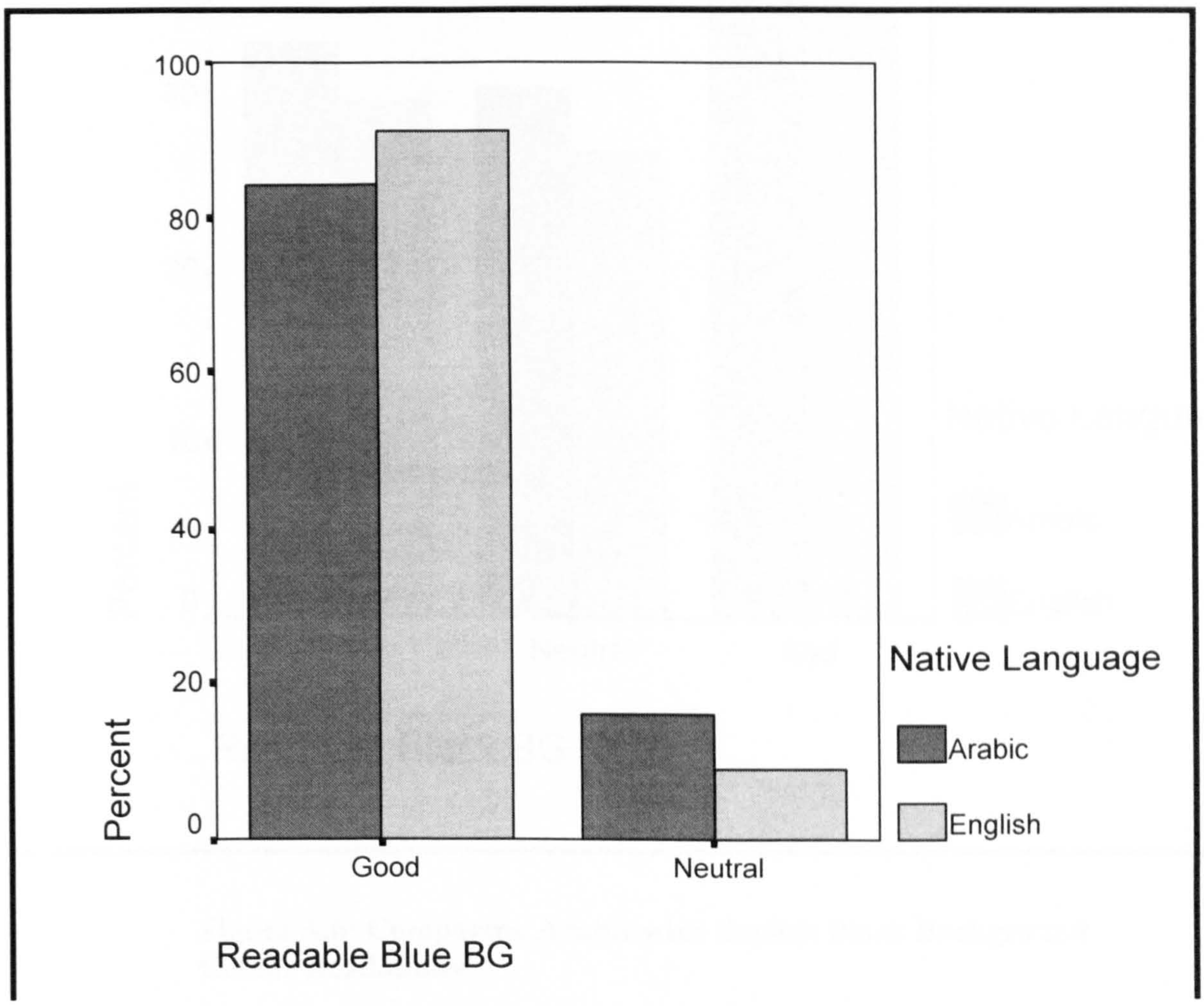


Figure A.5: Comparing Arabic with English Blue Background Colour Readability

A.1.2.2 Black Background Readability

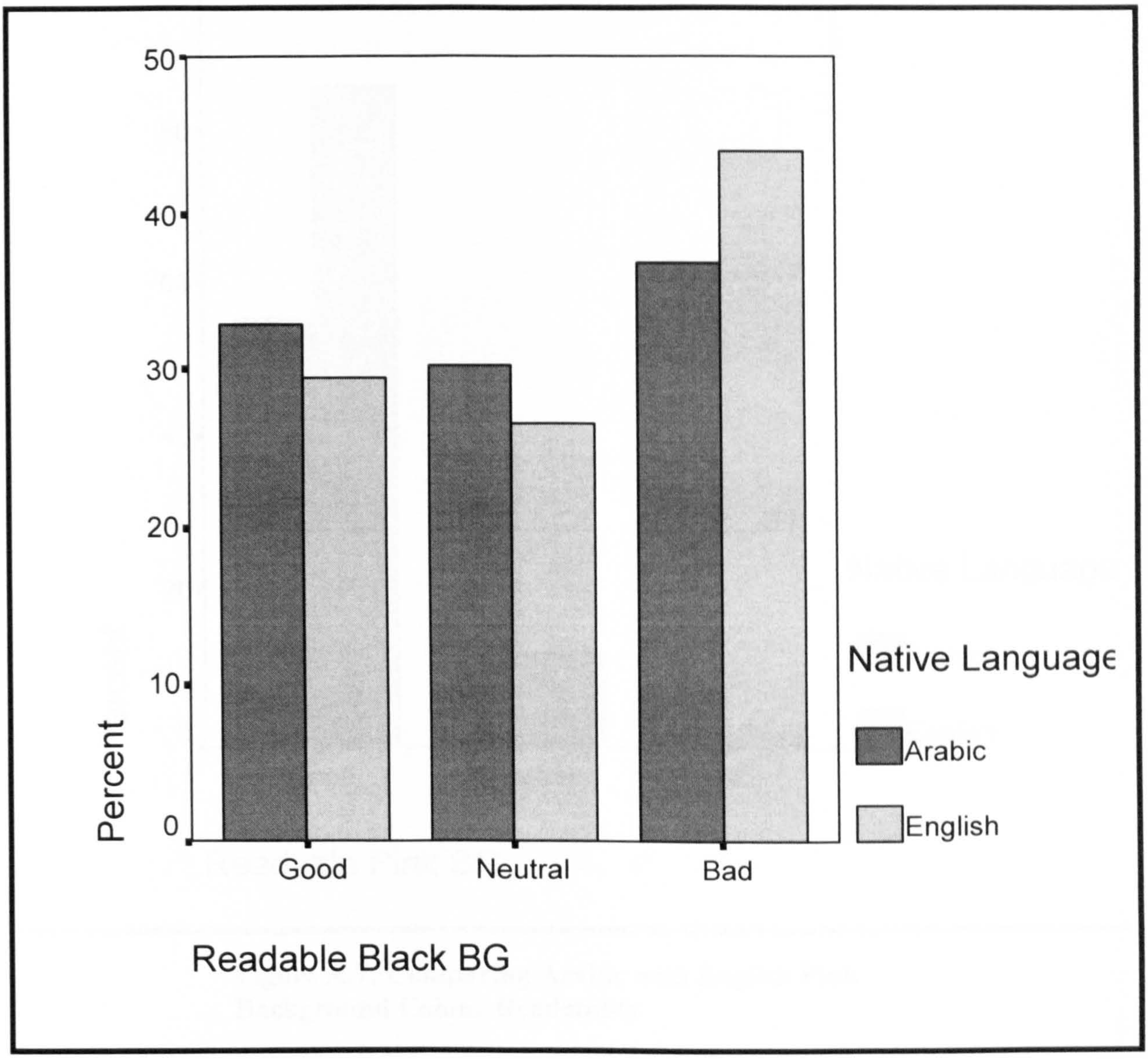


Figure A.6: Comparing Arabic with English Black Background Colour Readability

A.1.2.3 Pink Background Readability

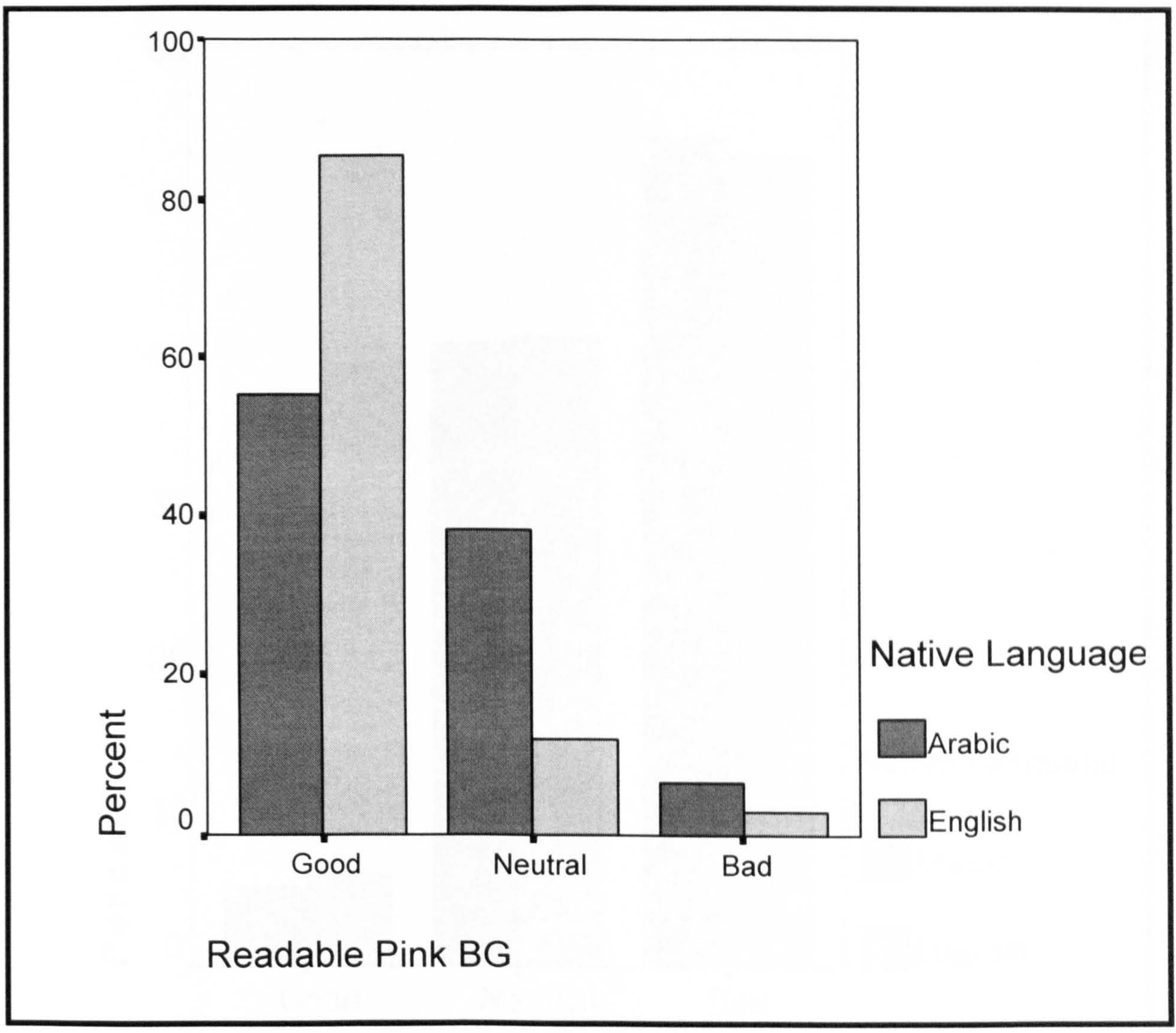


Figure A.7: Comparing Arabic with English Pink Background Colour Readability

A.1.2.4 Green Background Readability

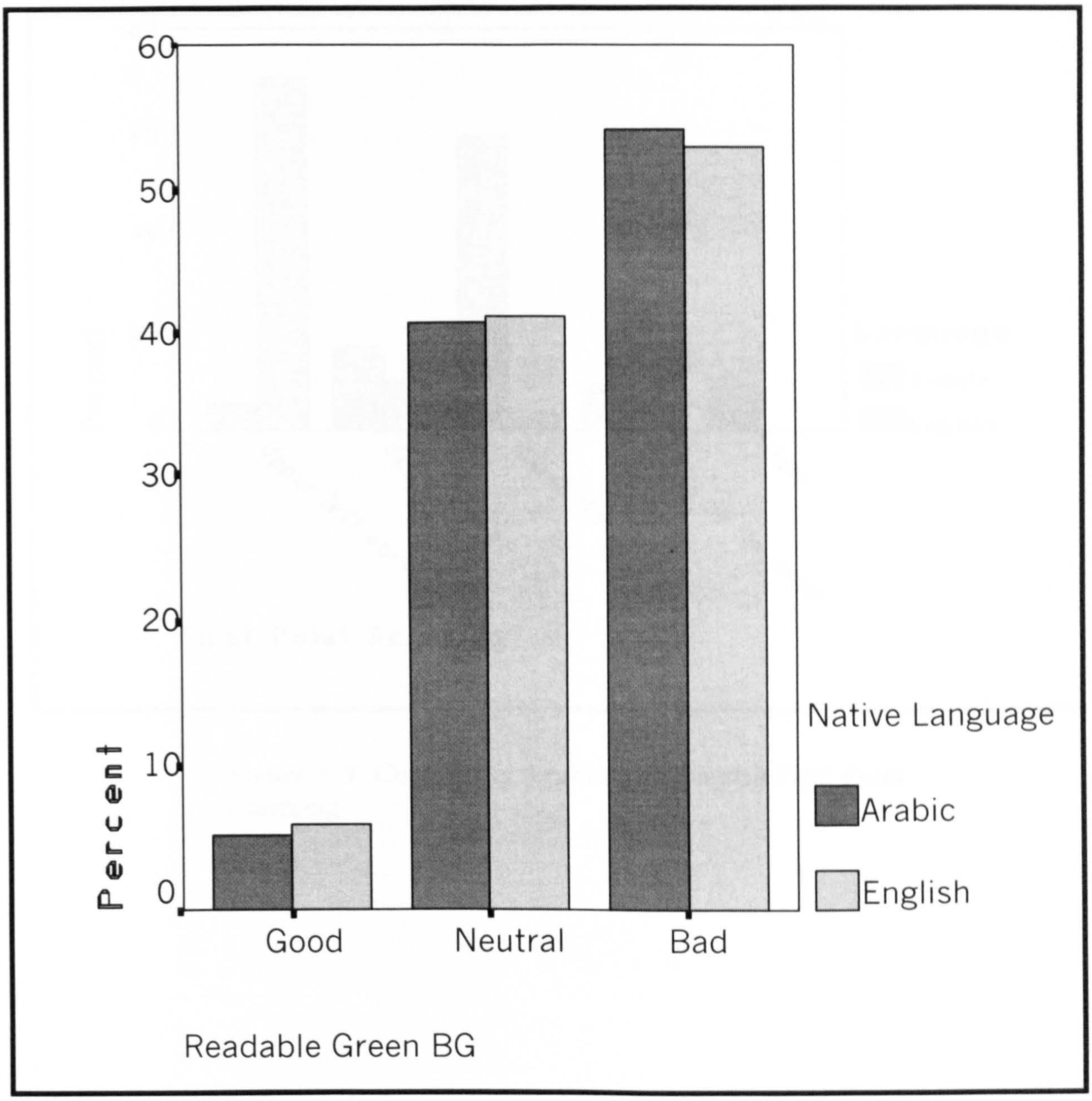


Figure A.8: Comparing Arabic with English Green Background Colour Readability

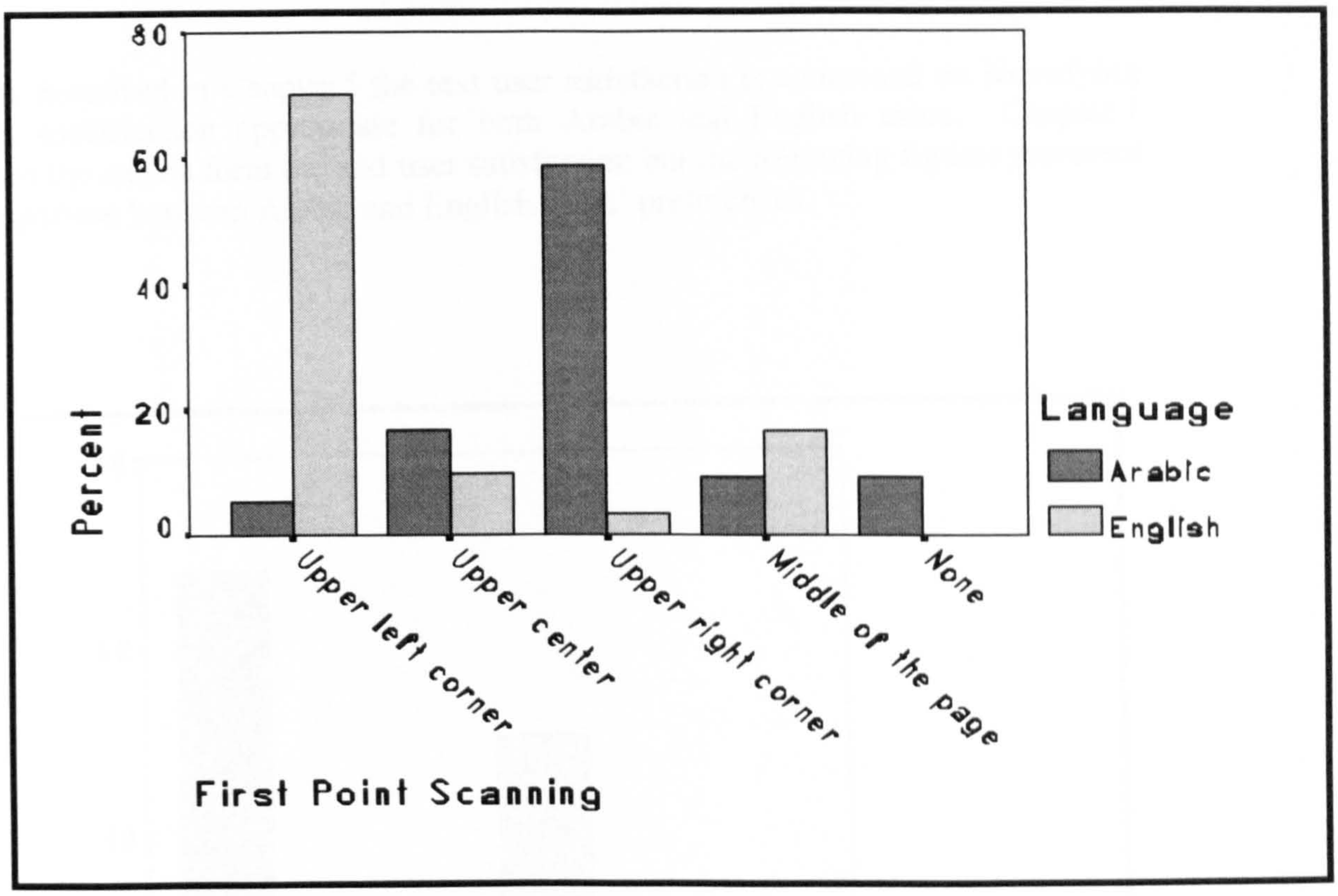


Figure A.9: Comparing Arabic with English First Point Scanning

A.2 Text User Satisfaction Findings Comparisons

As described in Chapter 5 the text user satisfaction is concerned on identifying the text specification appropriate for both Arabic and English users. Chapter 7 discussed the results form the text user satisfaction but the following figures represent the comparison between Arabic and English users' preferences.

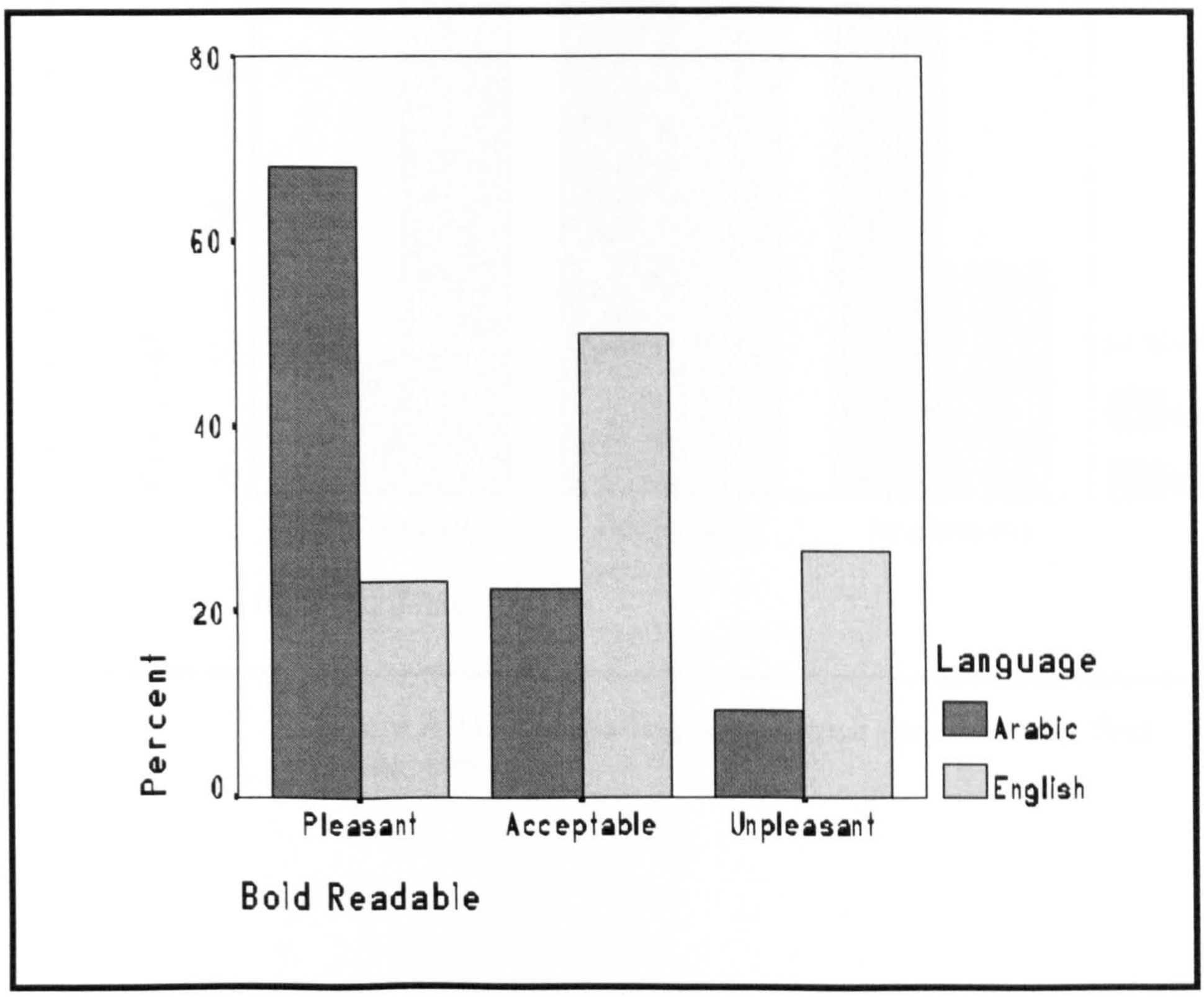


Figure A.10: Comparing Arabic with English Bold Text Readability

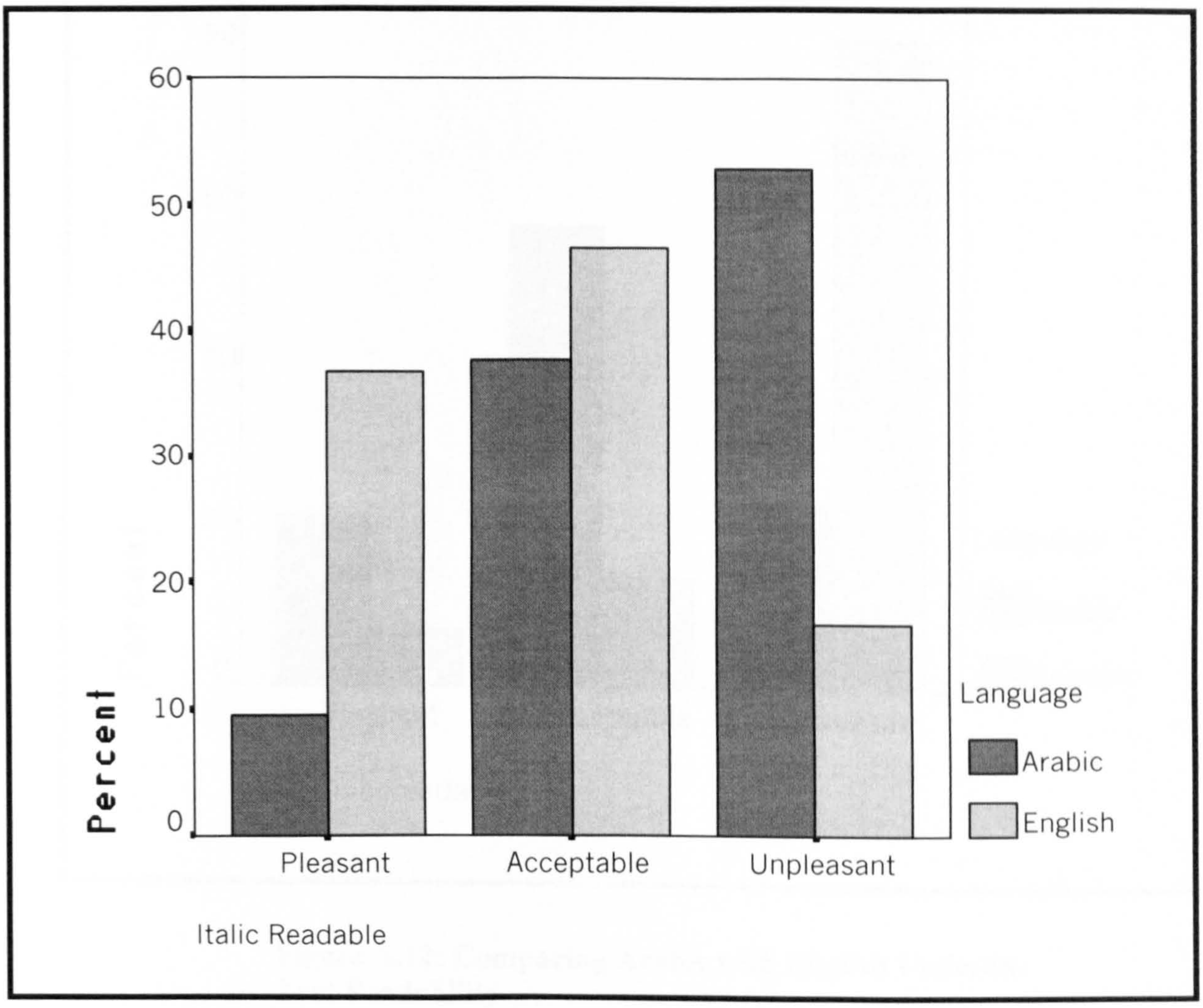


Figure A.11: Comparing Arabic with English *Italic* Text Readability

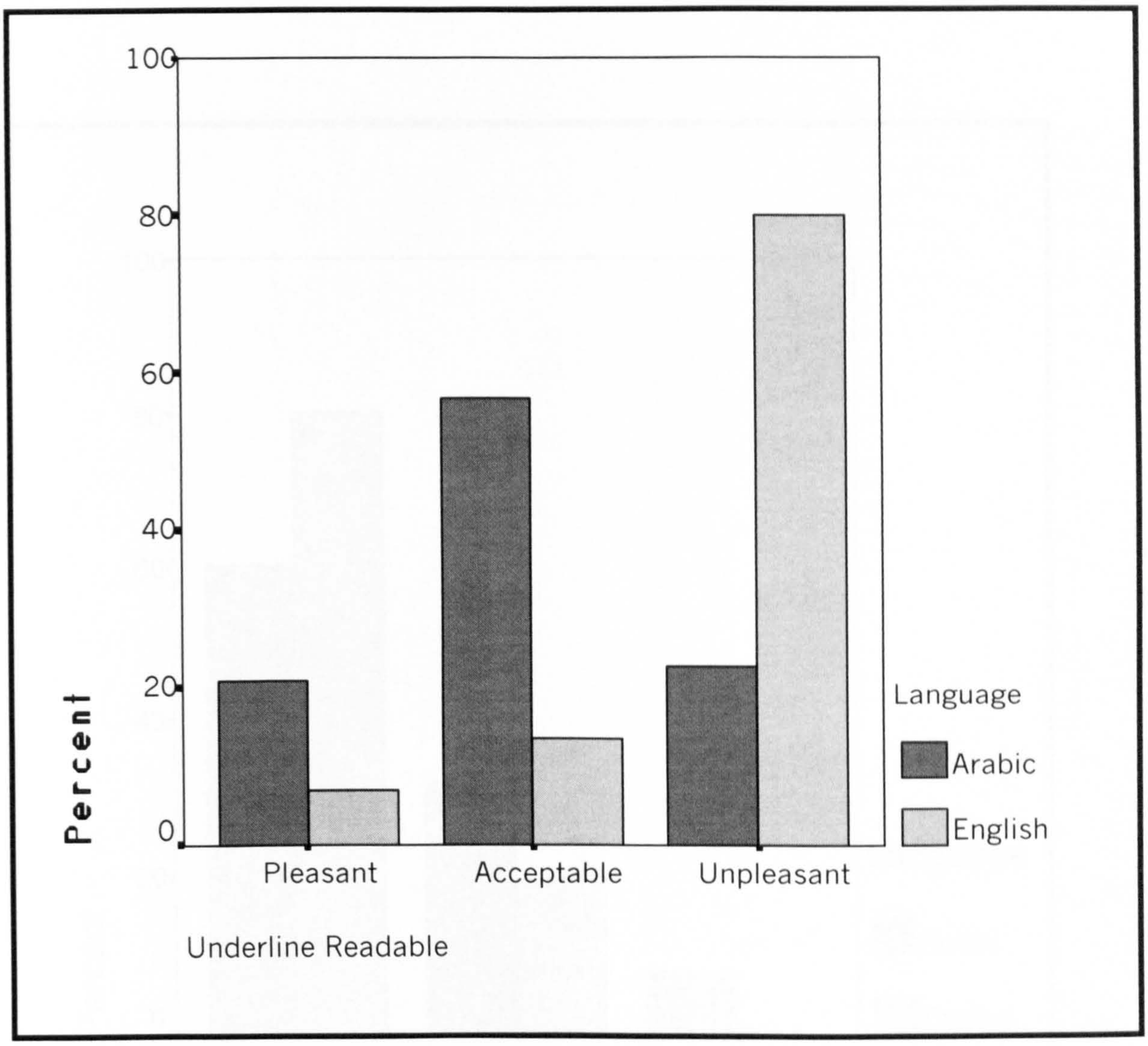


Figure A.12: Comparing Arabic with English Underline Text Readability

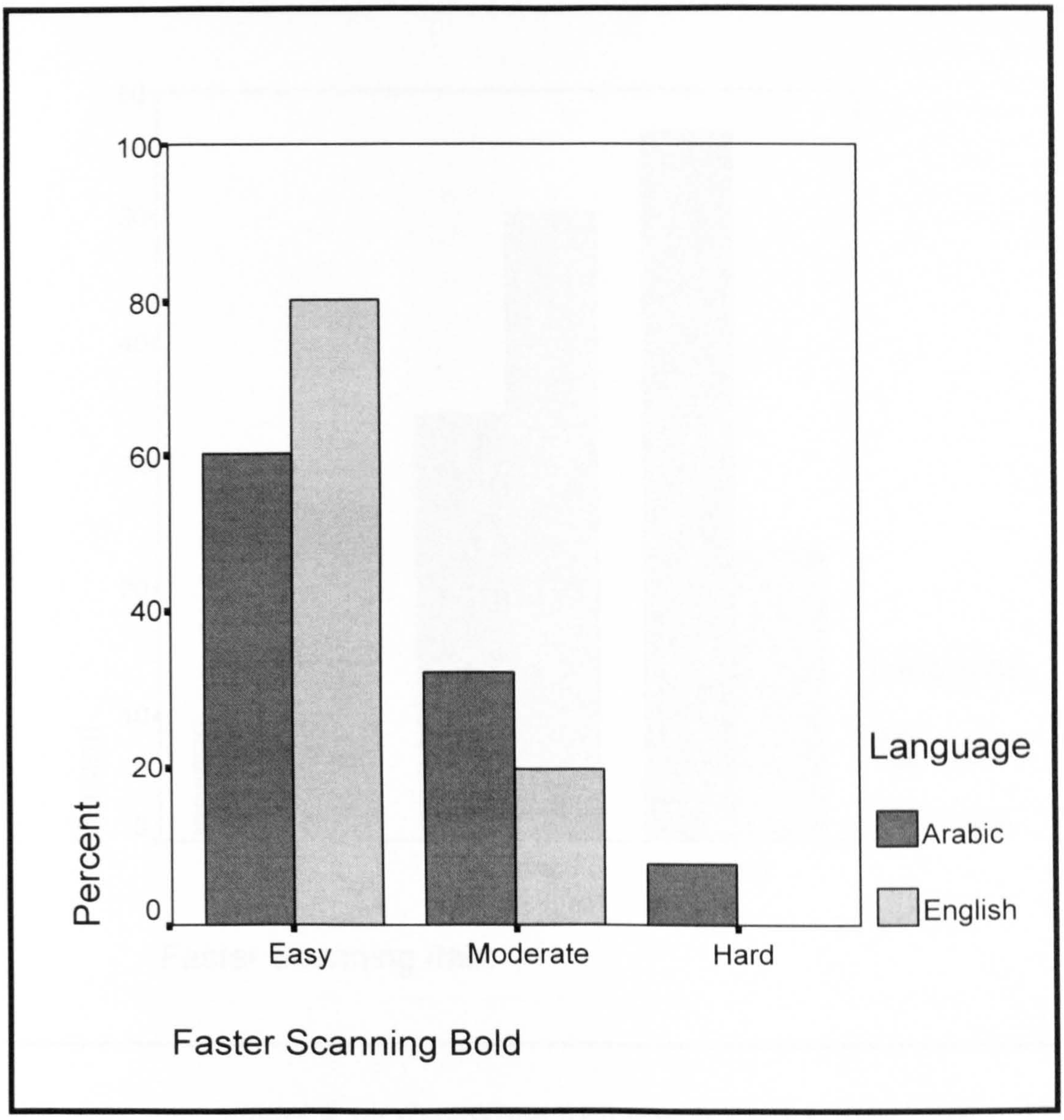


Figure A.13: Comparing Arabic with English Bold Text Scanability

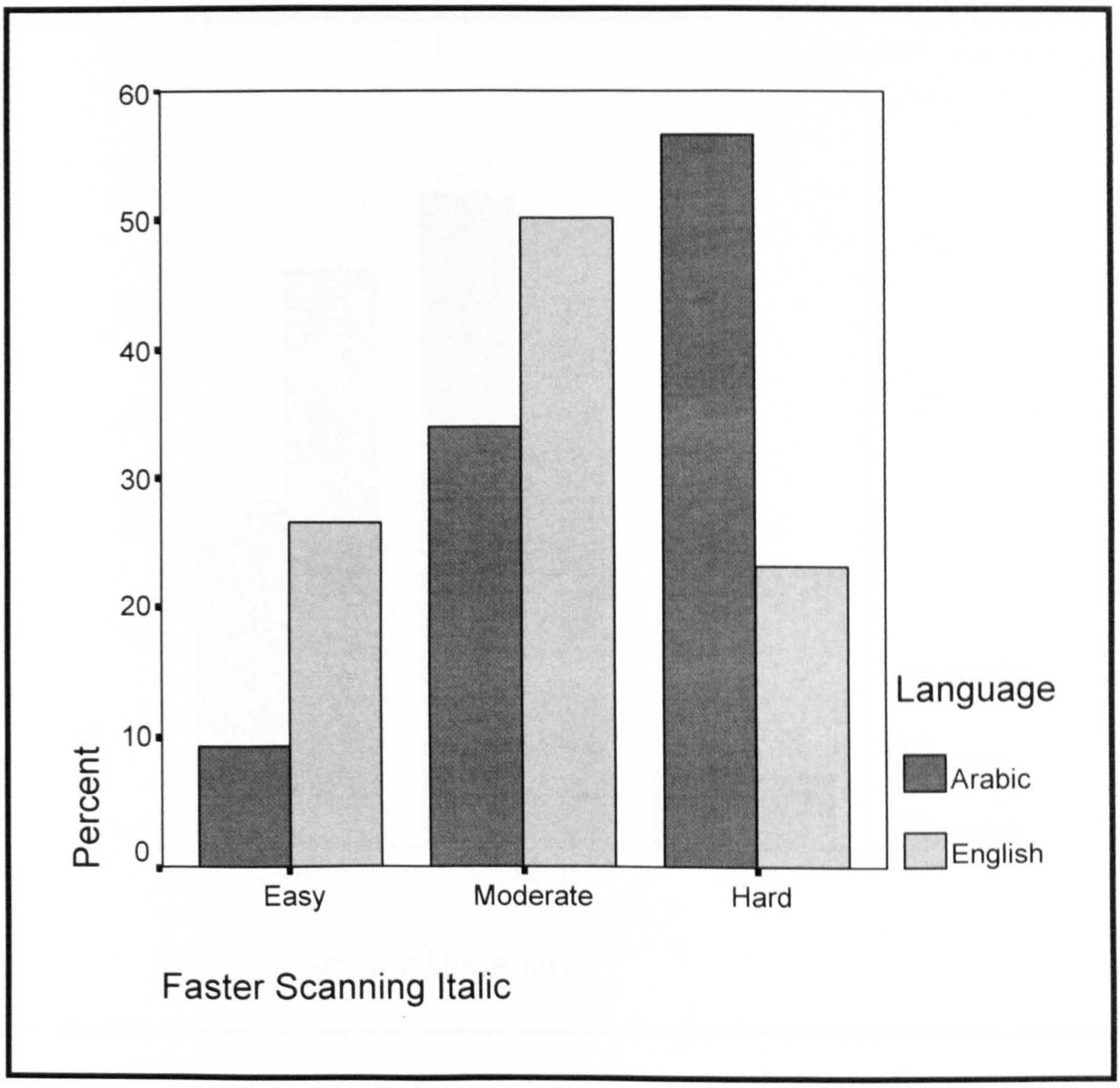


Figure A.14: Comparing Arabic with English Italic Text Scanability

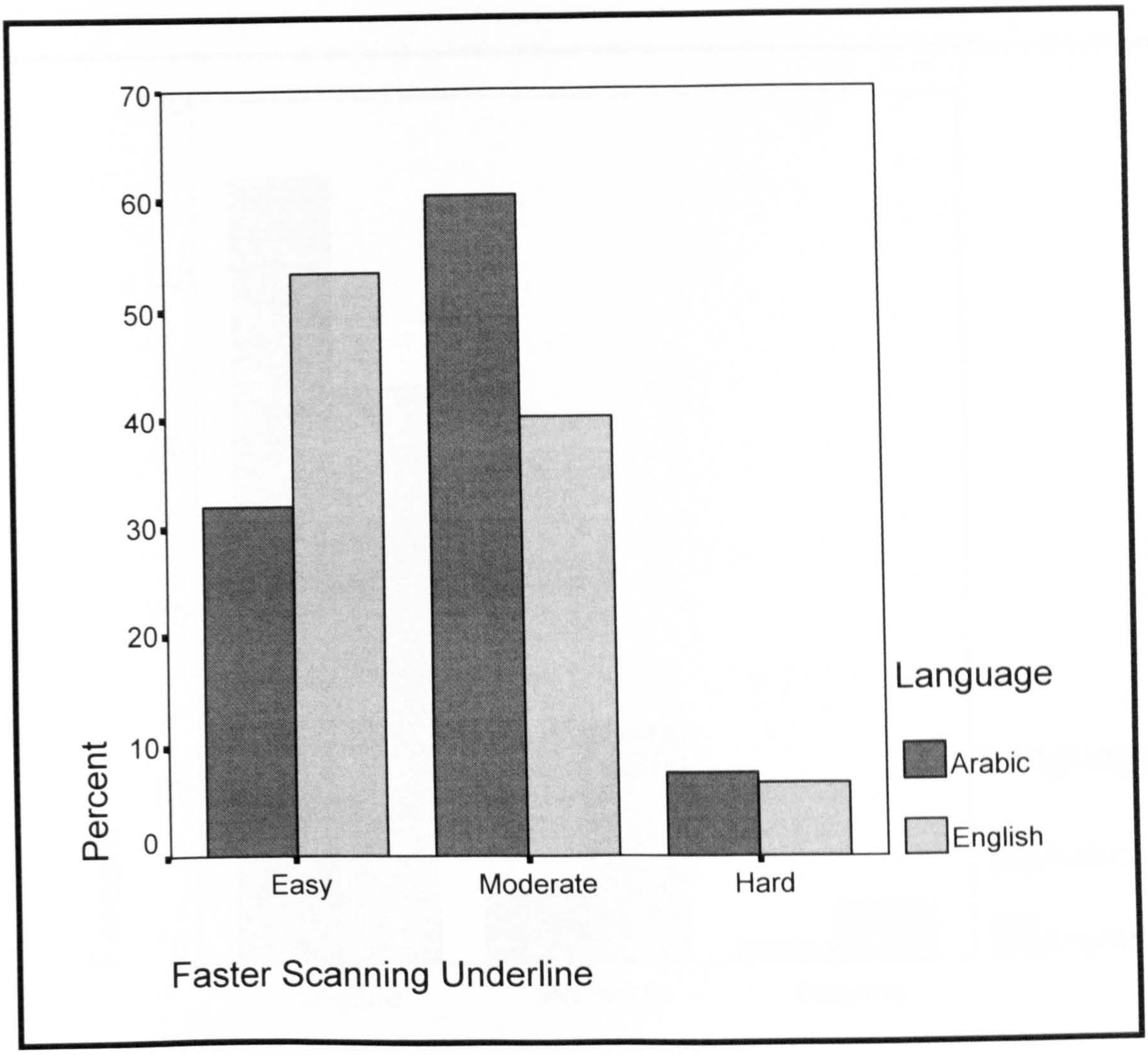


Figure A.15: Comparing Arabic with English Underline Text Scanability

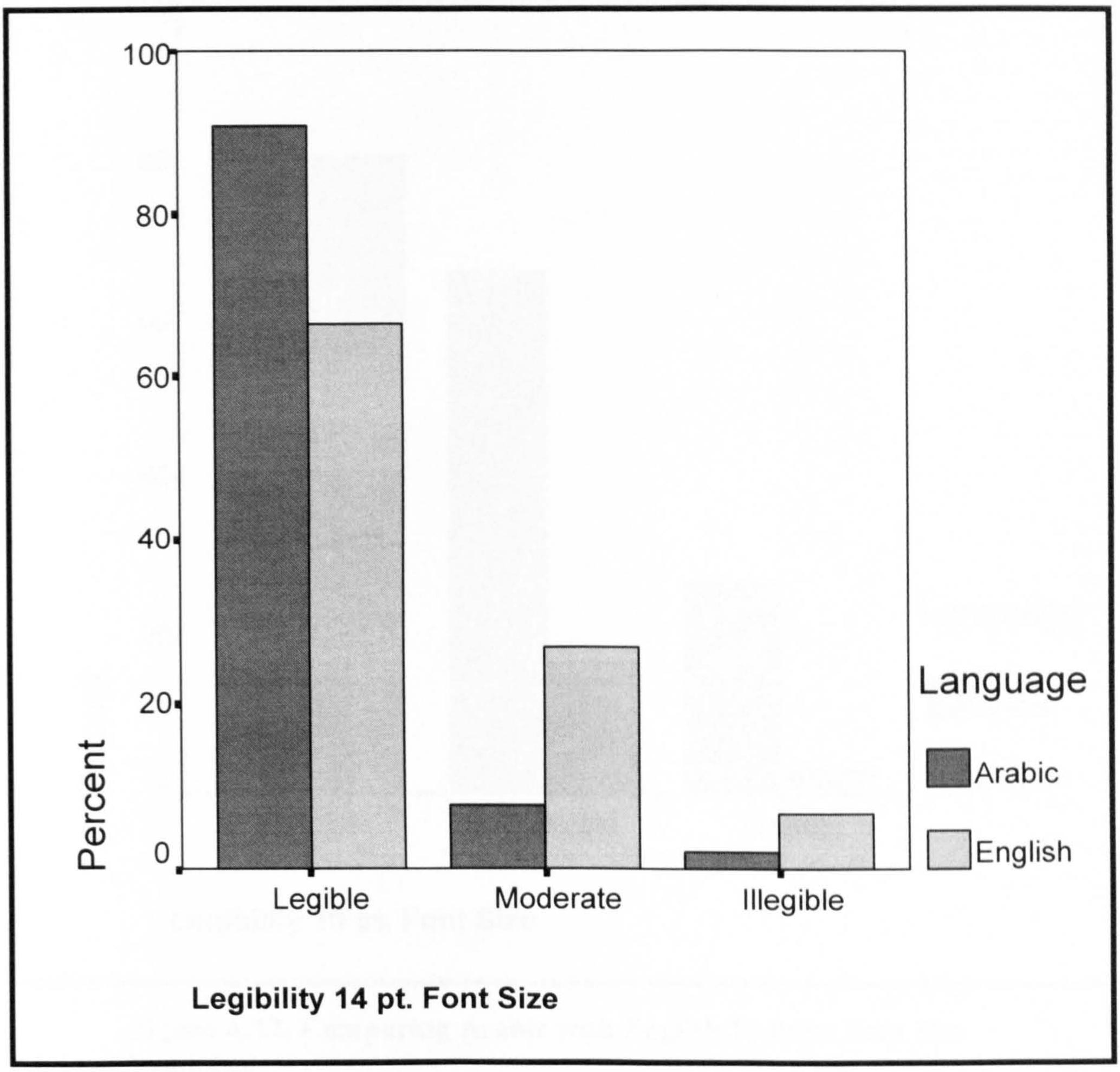


Figure A.16: Comparing Arabic with English 14 Points Font Size Legibility

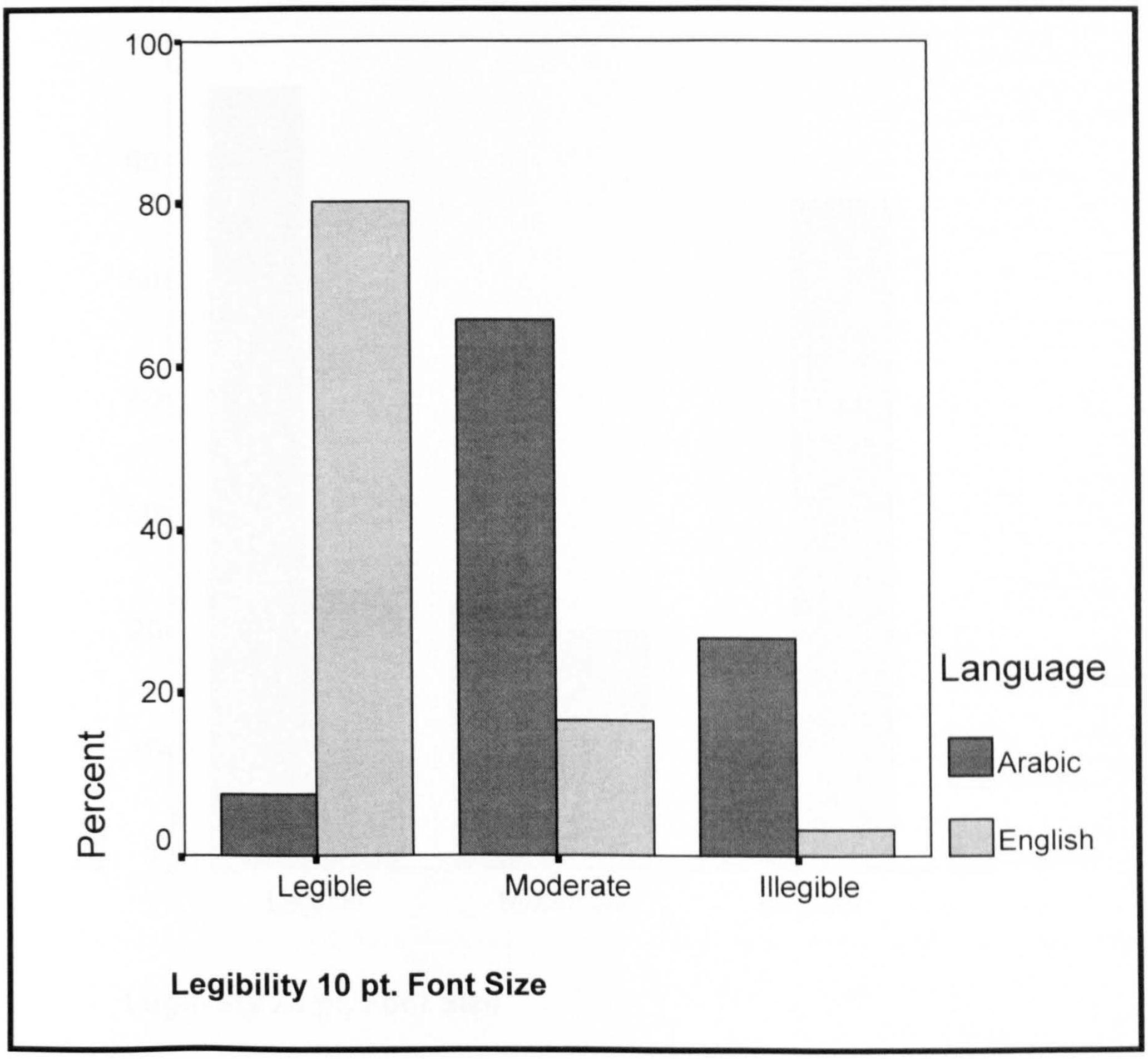


Figure A.17: Comparing Arabic with English 10 Point Font Size Legibility

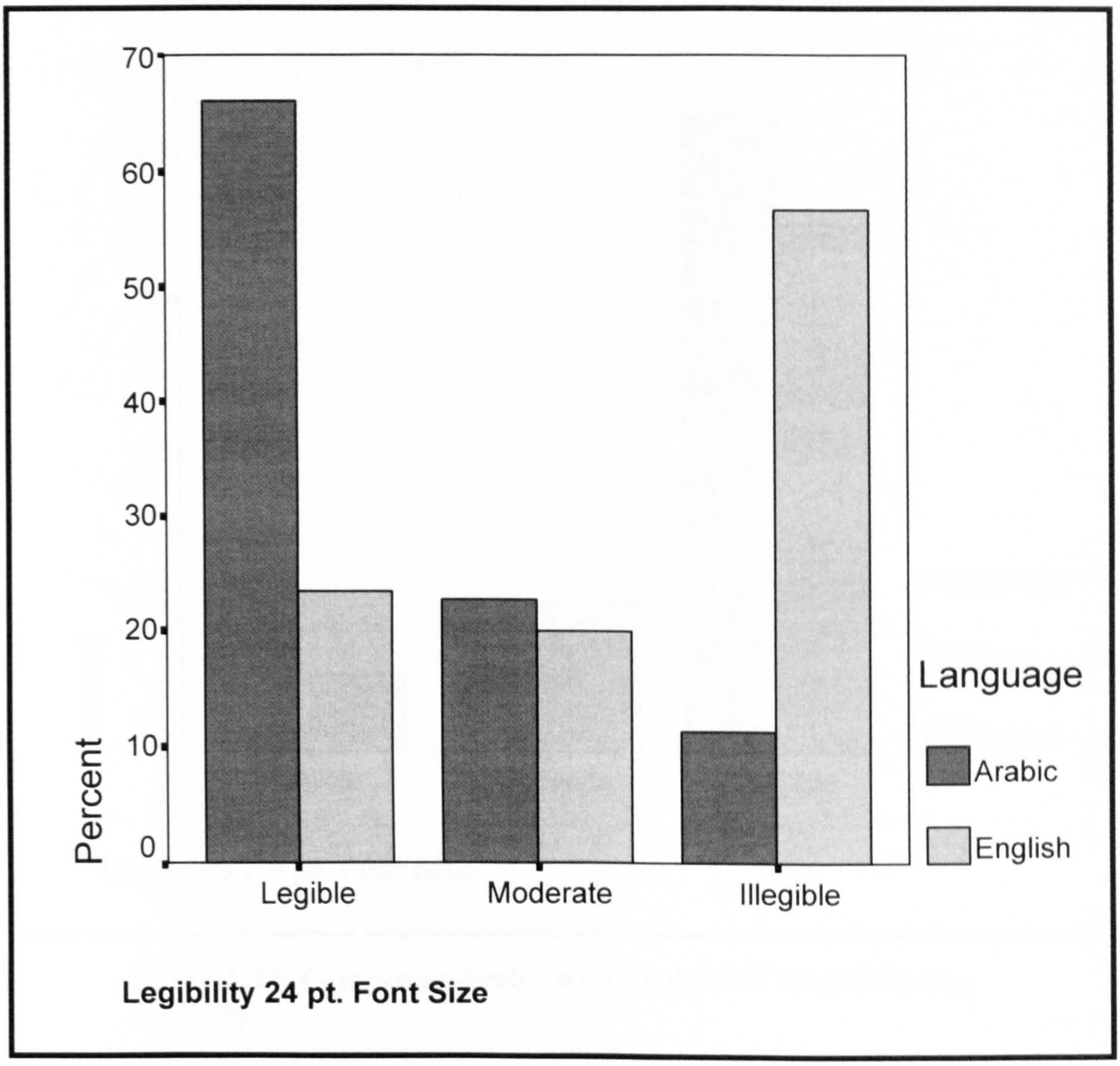


Figure A.18: Comparing Arabic with English 24 Point Font Size Legibility

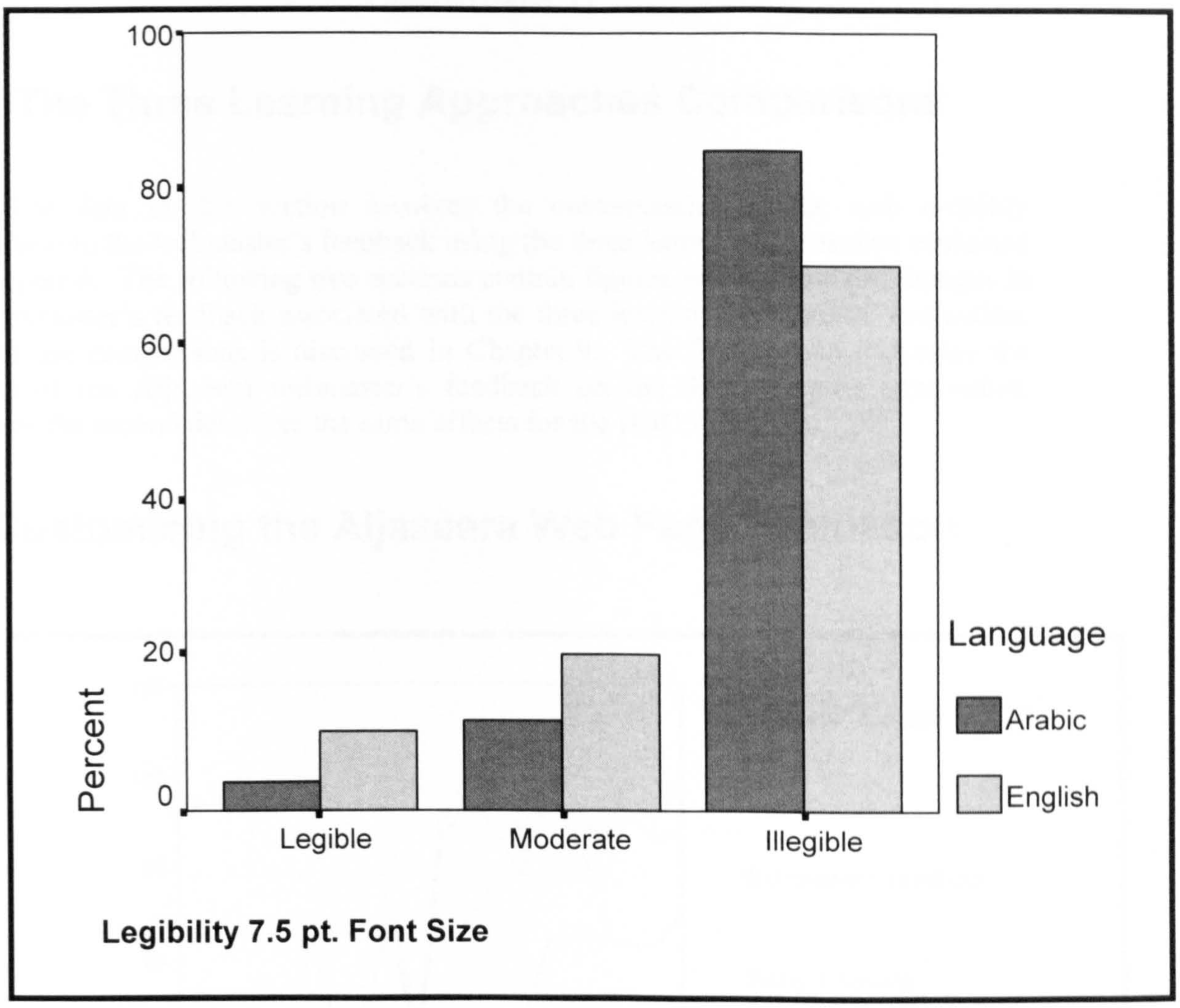


Figure A.19: Comparing Arabic with English 7.5 Point Font Size Legibility

Appendix B

The Three Learning Approaches Comparisons

The data in this section involves the customisation in the web usability evaluation to the webmaster's feedback using the three learning approaches explained in Chapter 6. The following two sections contain figures which show the changes in the webmaster's feedback associated with the three learning approaches' evaluation, part of the comparisons is discussed in Chapter 9. The first section illustrates the effects of the Aljazeera webmaster's feedback on the three learning approaches, whereas the second describes the same effects for the BBC web page.

B.1 Customising the Aljazeera Web Page Evaluation

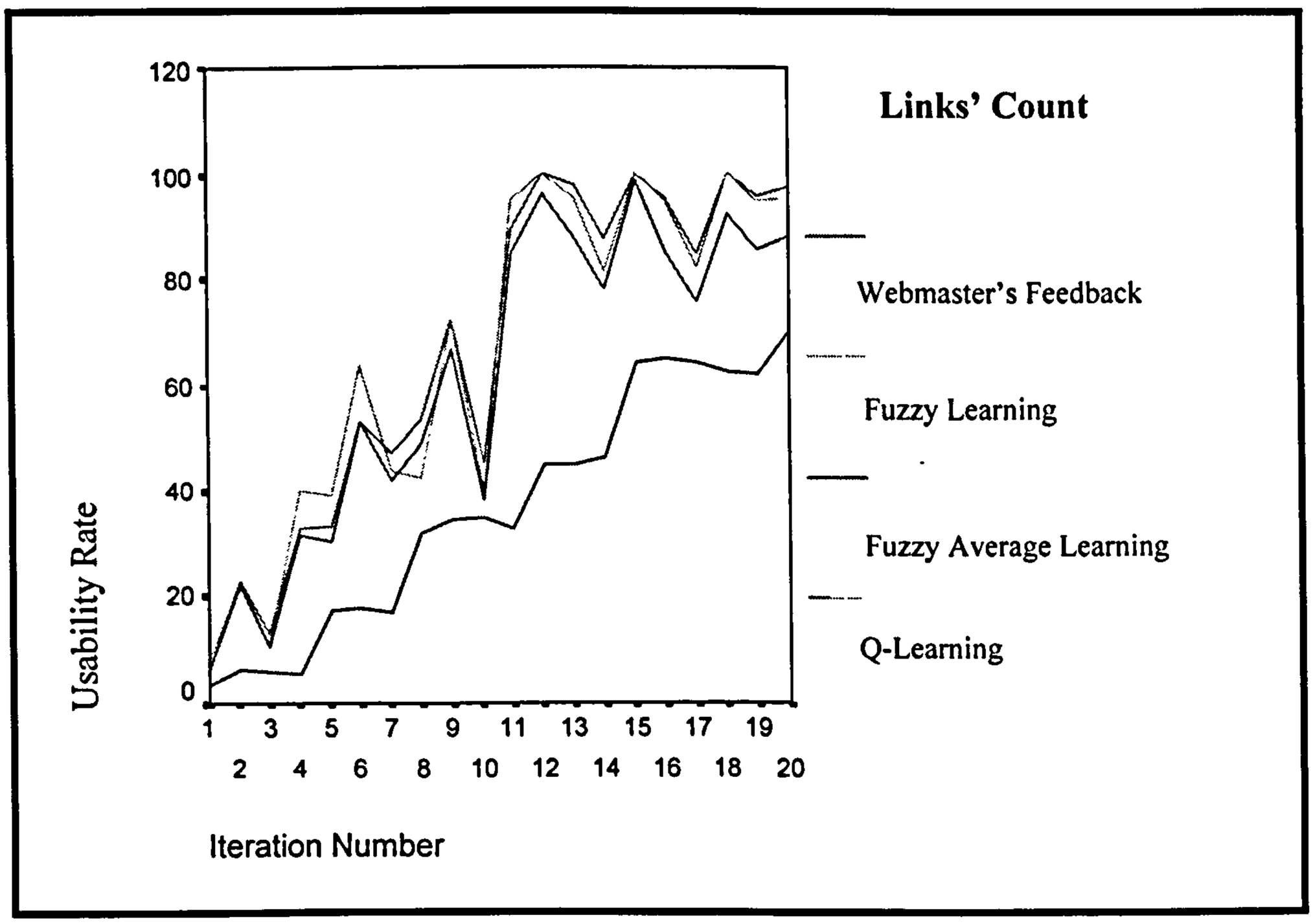


Figure B.1: The Customisation of the Number of Links Count Usability Rate in Aljazeera Web Page

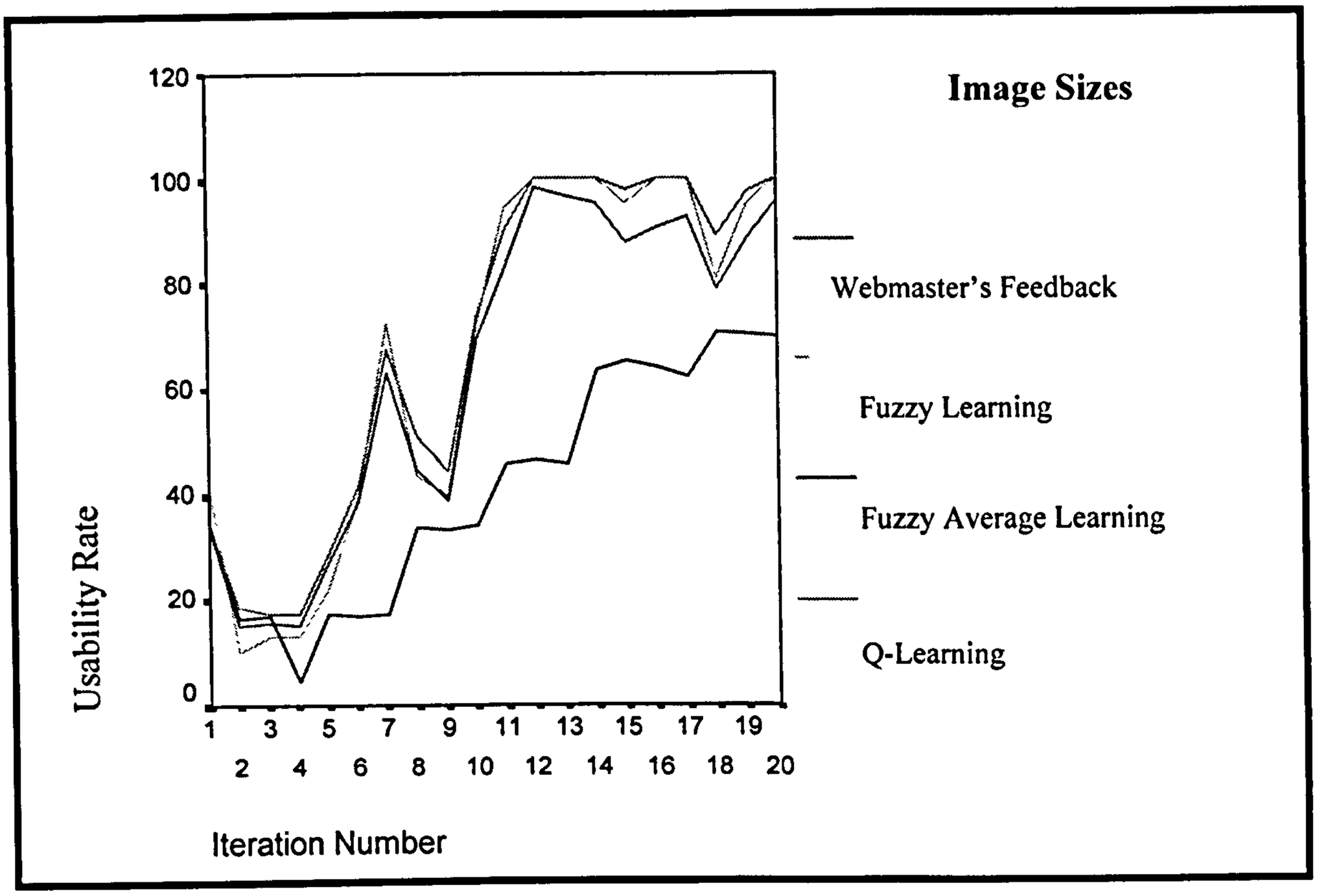


Figure B.2: The Customisation of the Image Sizes Usability Rate in Aljazeera Web Page

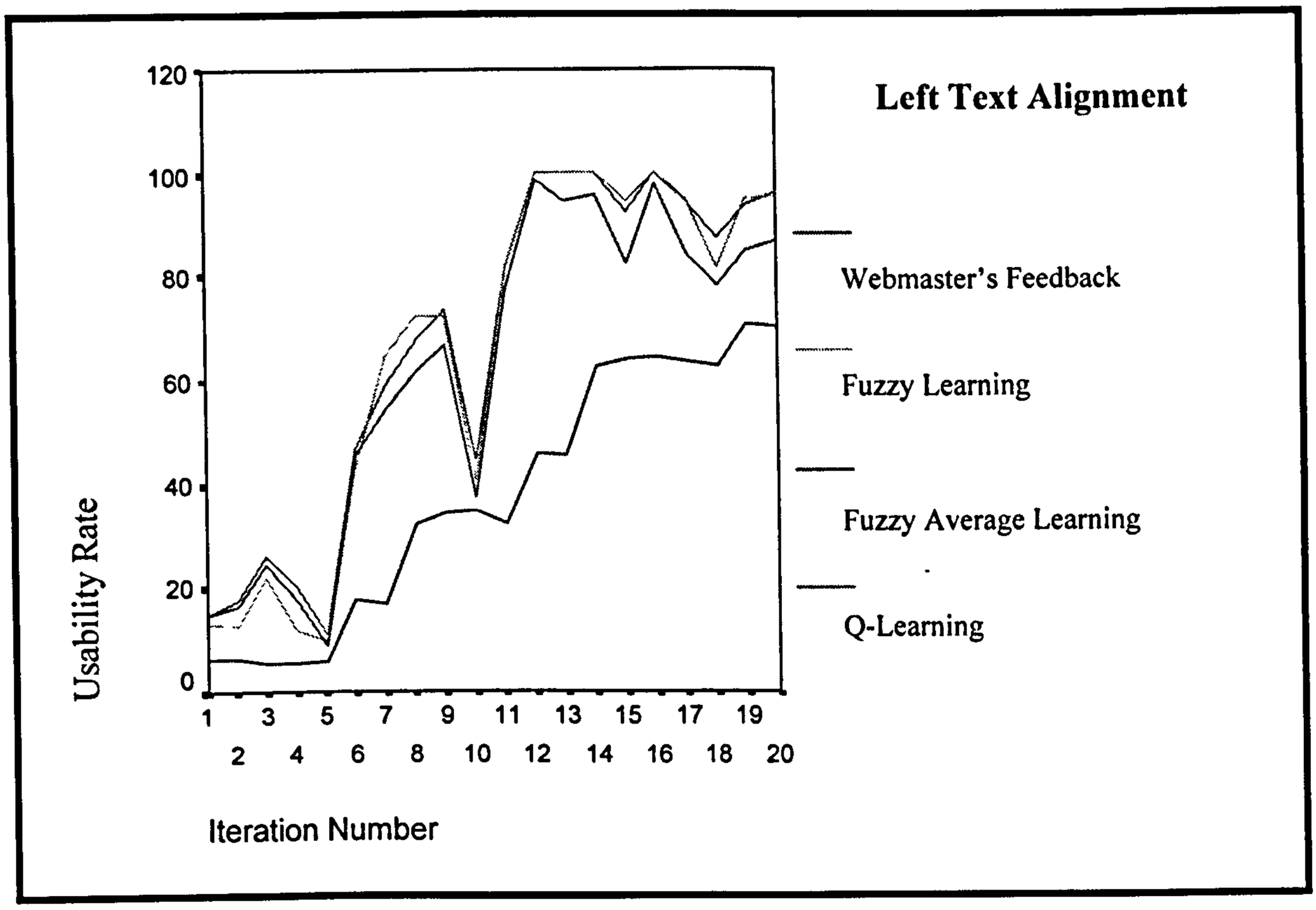


Figure B.3: The Customisation of the Left Text Alignment Usability Rate in Aljazeera Web Page

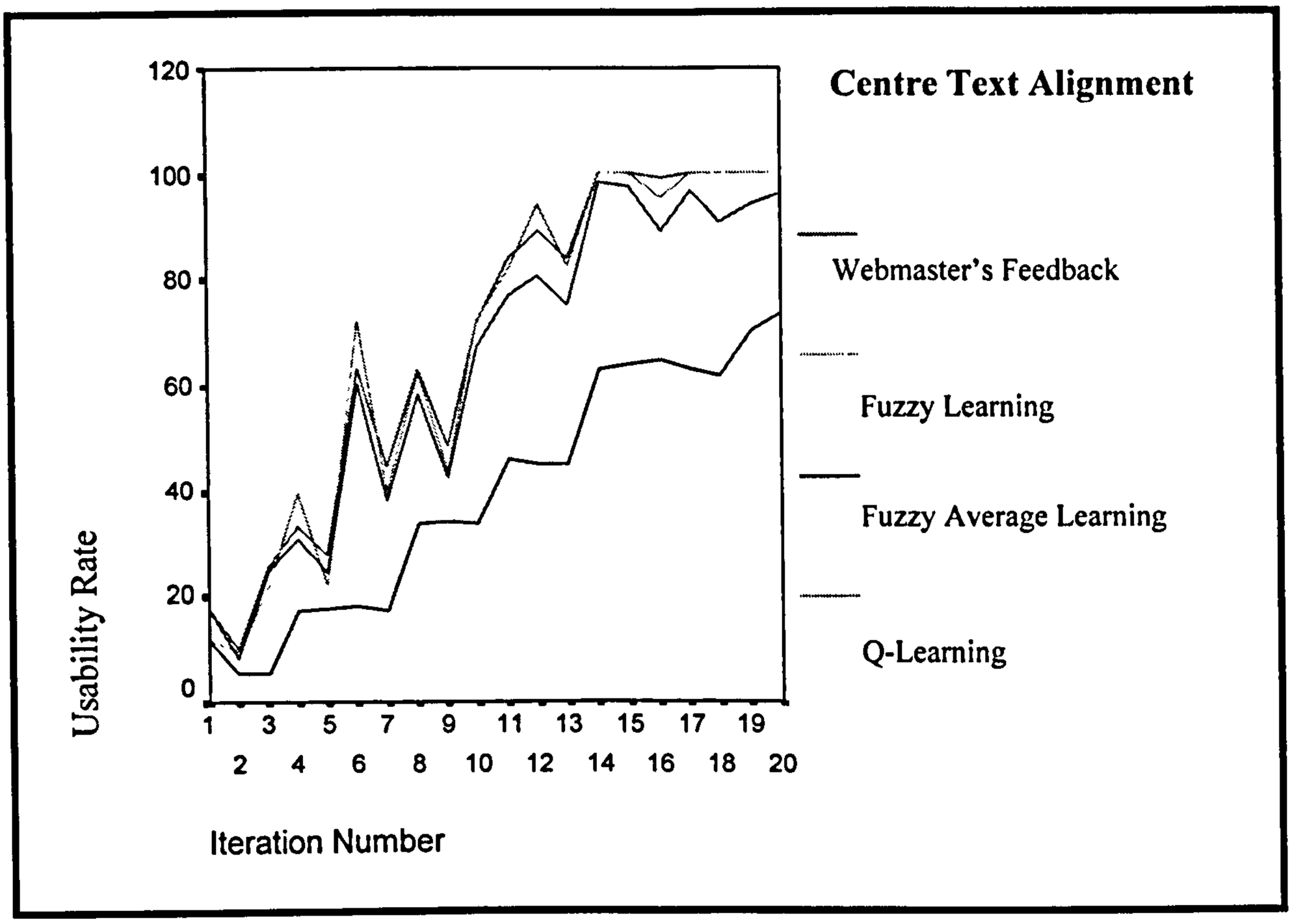


Figure B.4: The Customisation of the Centre Text Alignment Usability Rate in Aljazeera Web Page

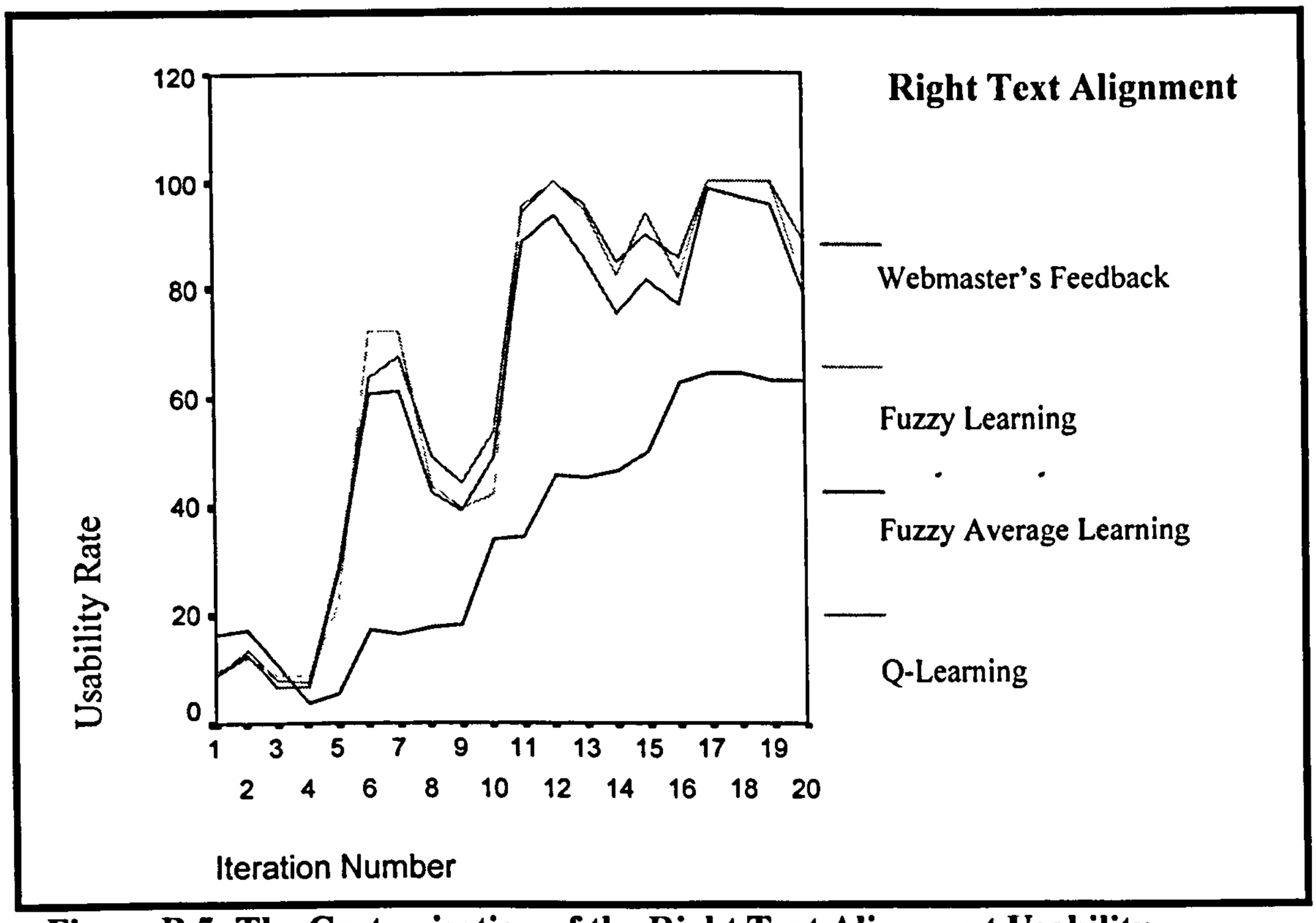


Figure B.5: The Customisation of the Right Text Alignment Usability Rate in Aljazeera Web Page

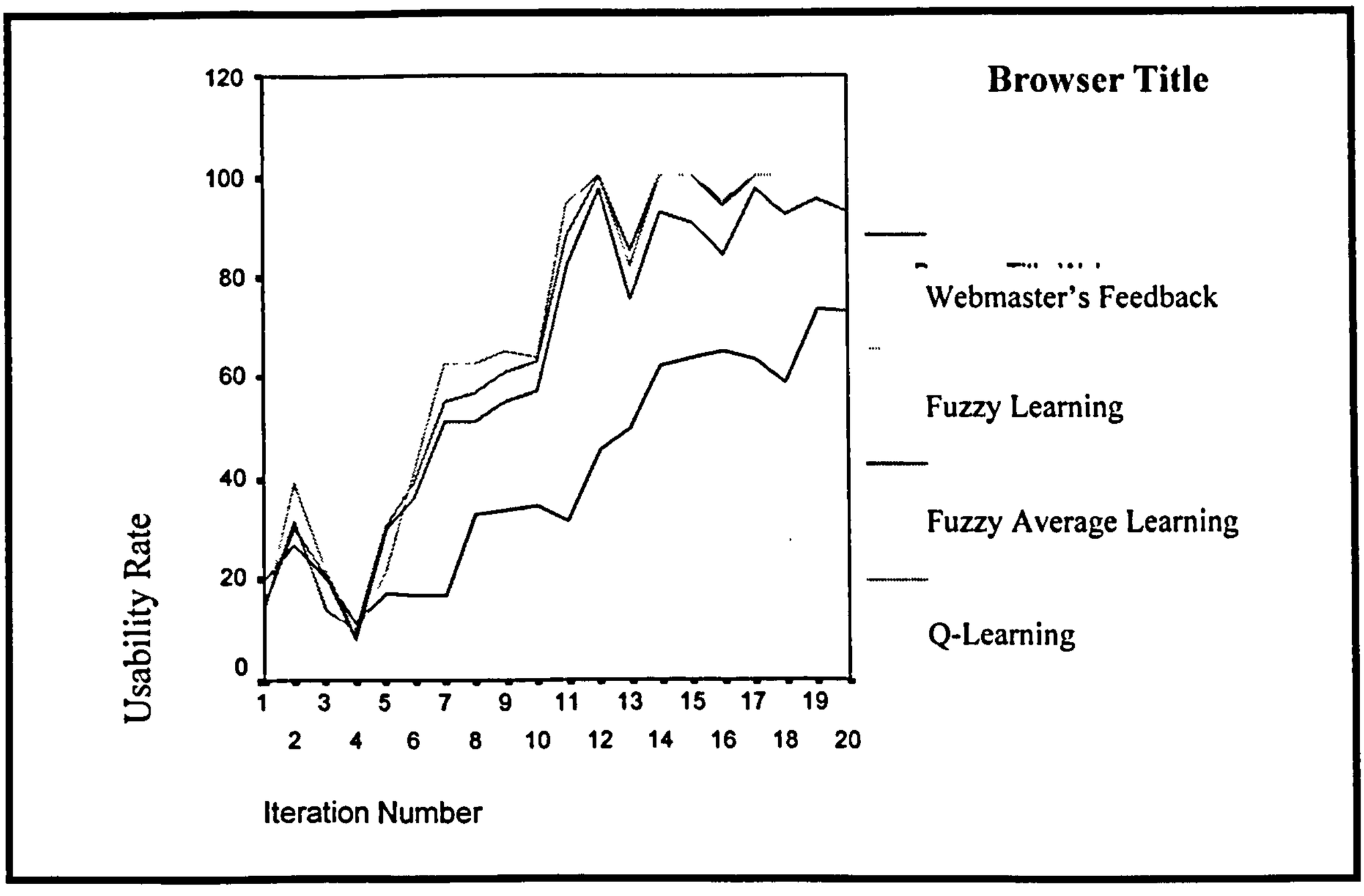


Figure B.6: The Customisation of the Browser Title Usability Rate in Aljazeera Web Page

B.2 Customising the BBC Web Page Evaluation

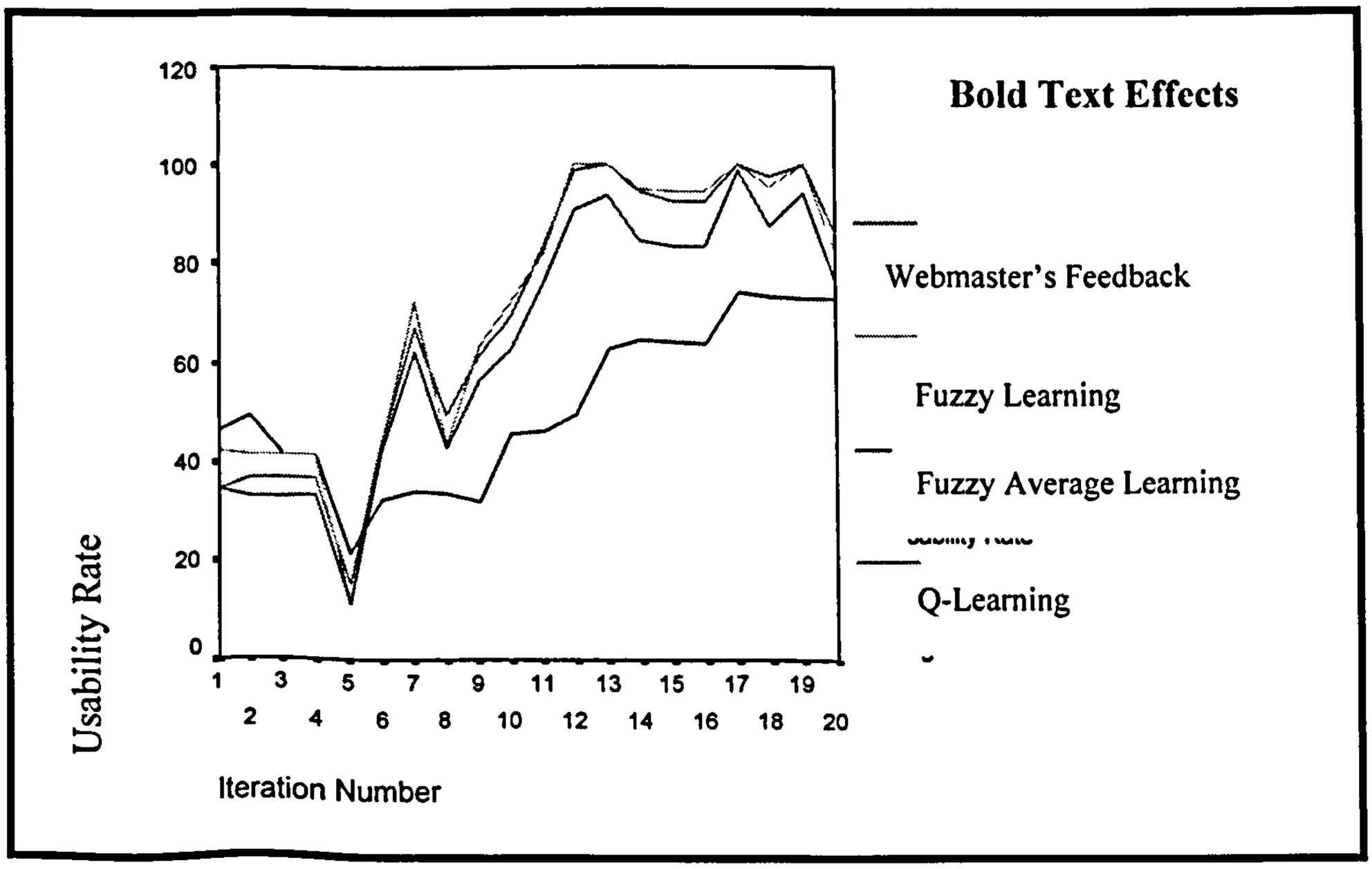


Figure B.7: The Customisation of Bold Text Effects Usability Rate in the BBC Web Page

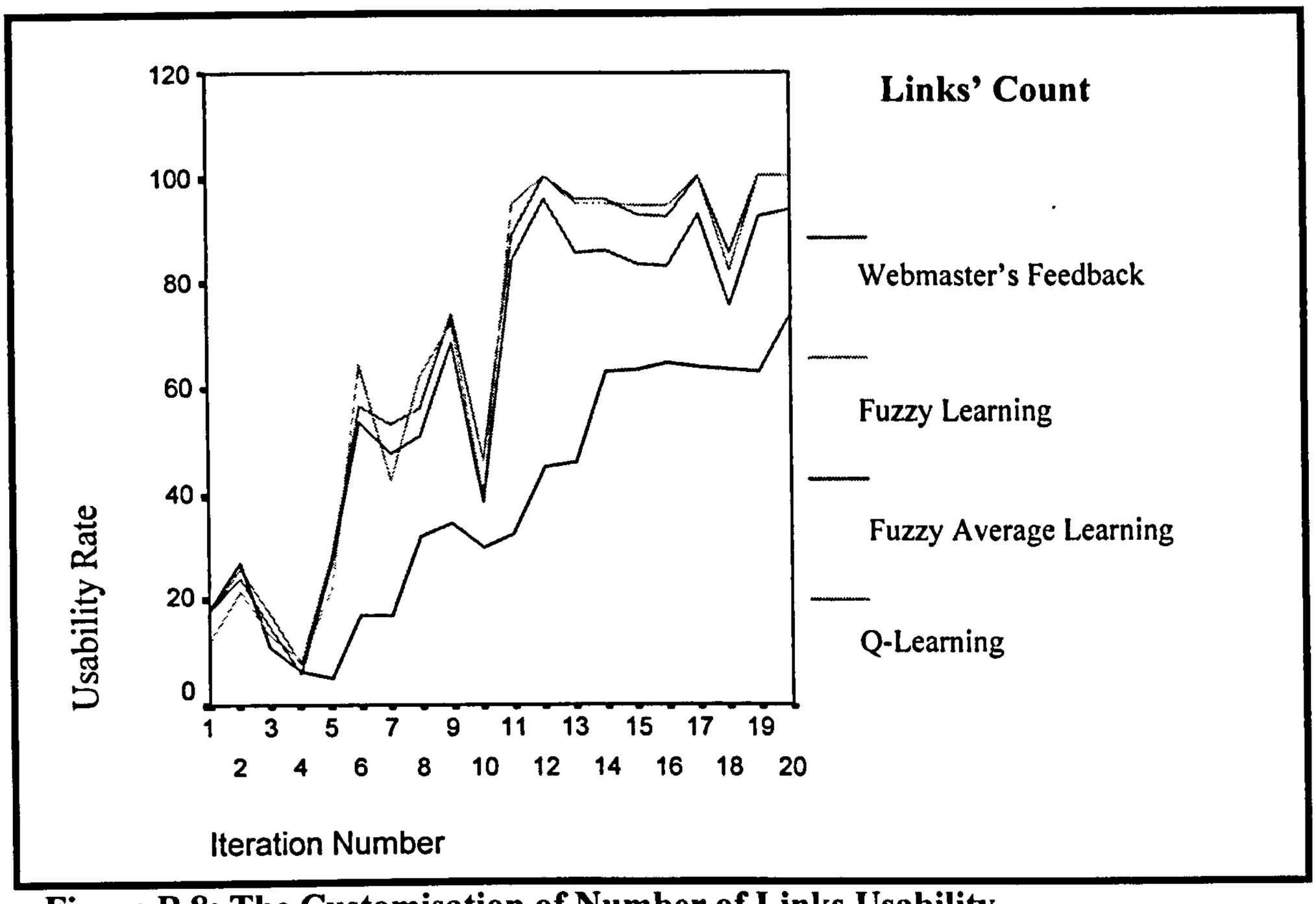


Figure B.8: The Customisation of Number of Links Usability Rate in the BBC Web Page

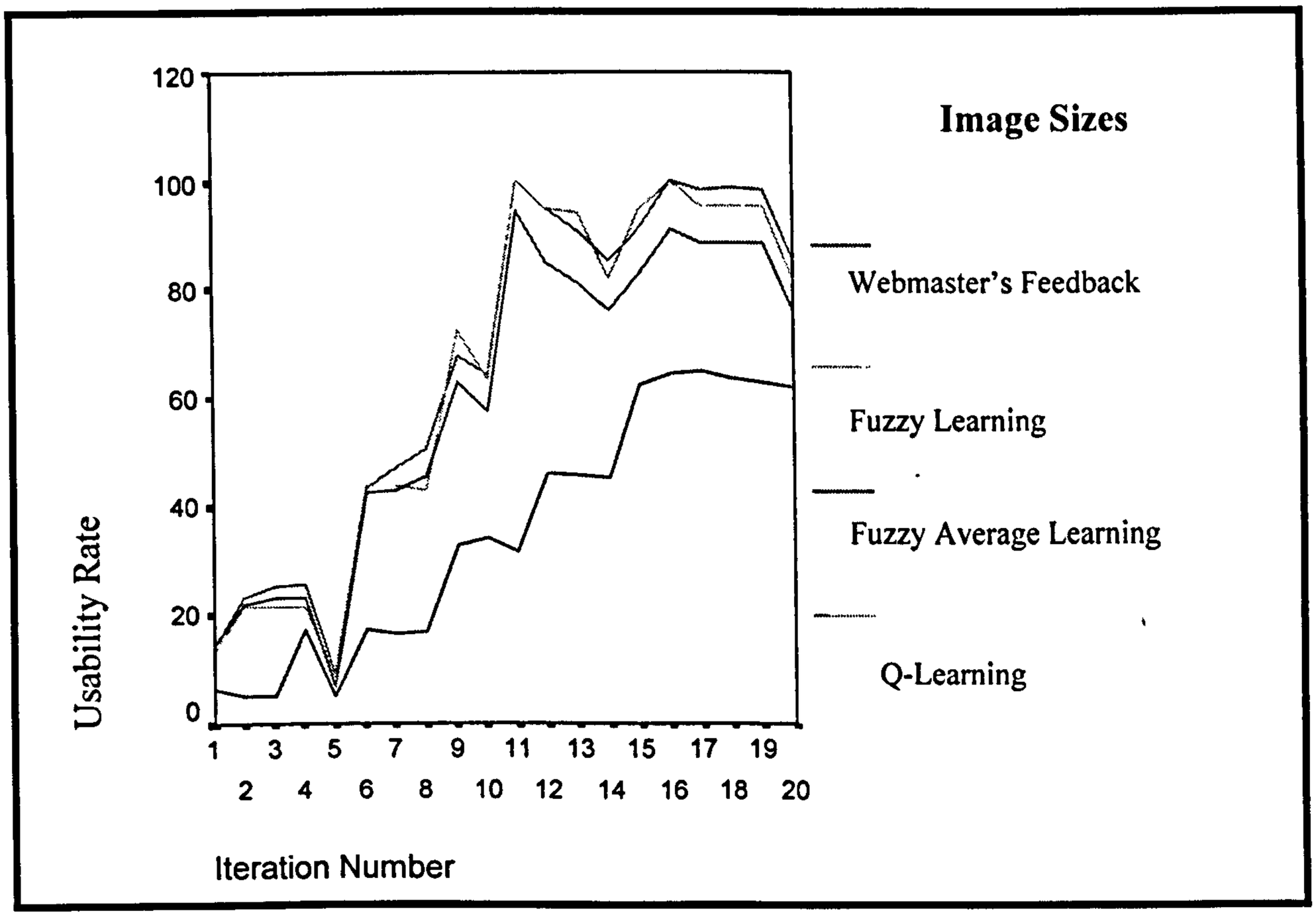


Figure B.9: The Customisation of the Image Sizes Usability Rate in the BBC Web Page

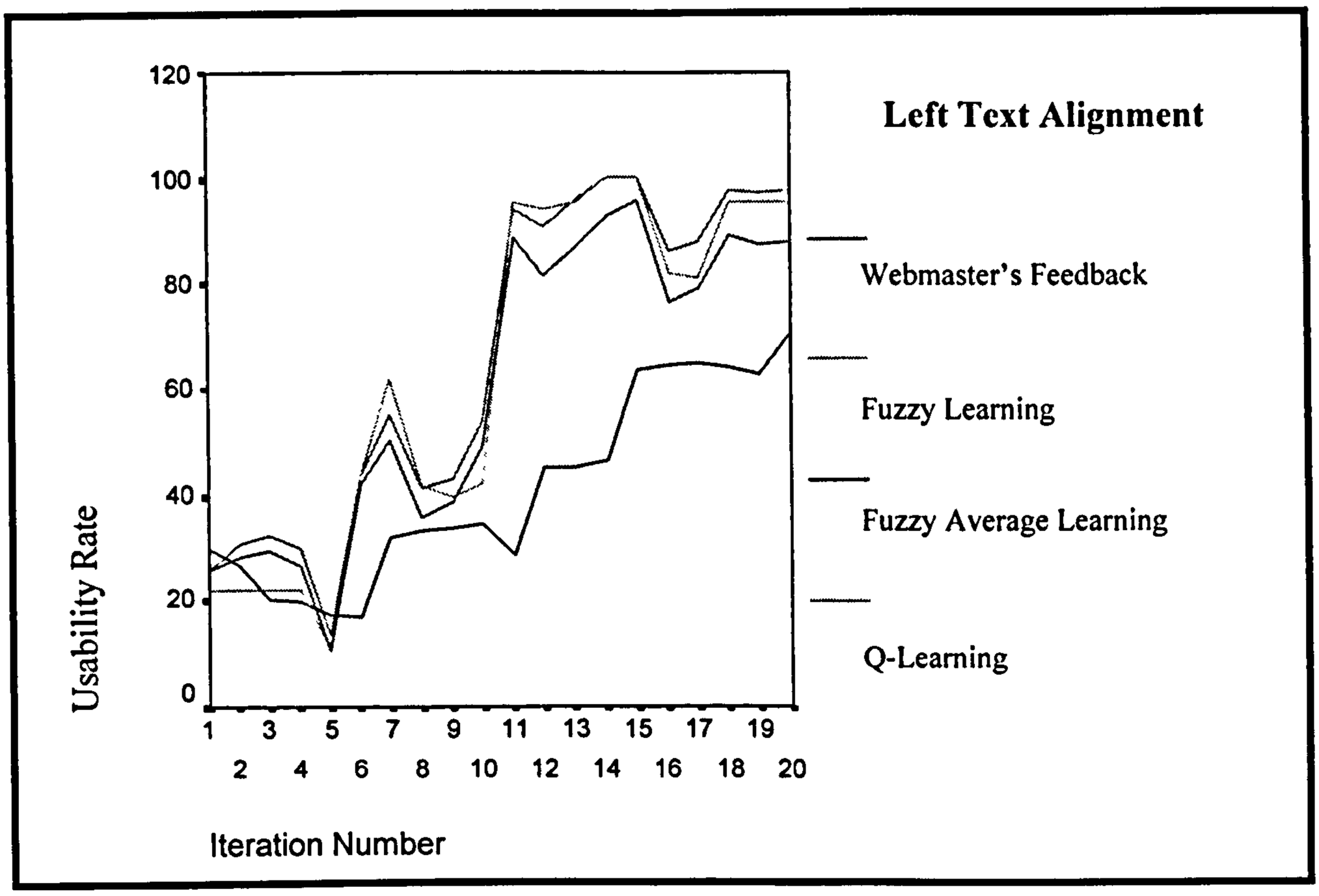


Figure B.10: The Customisation of the Left Text Alignment Usability Rate in the BBC Web Page

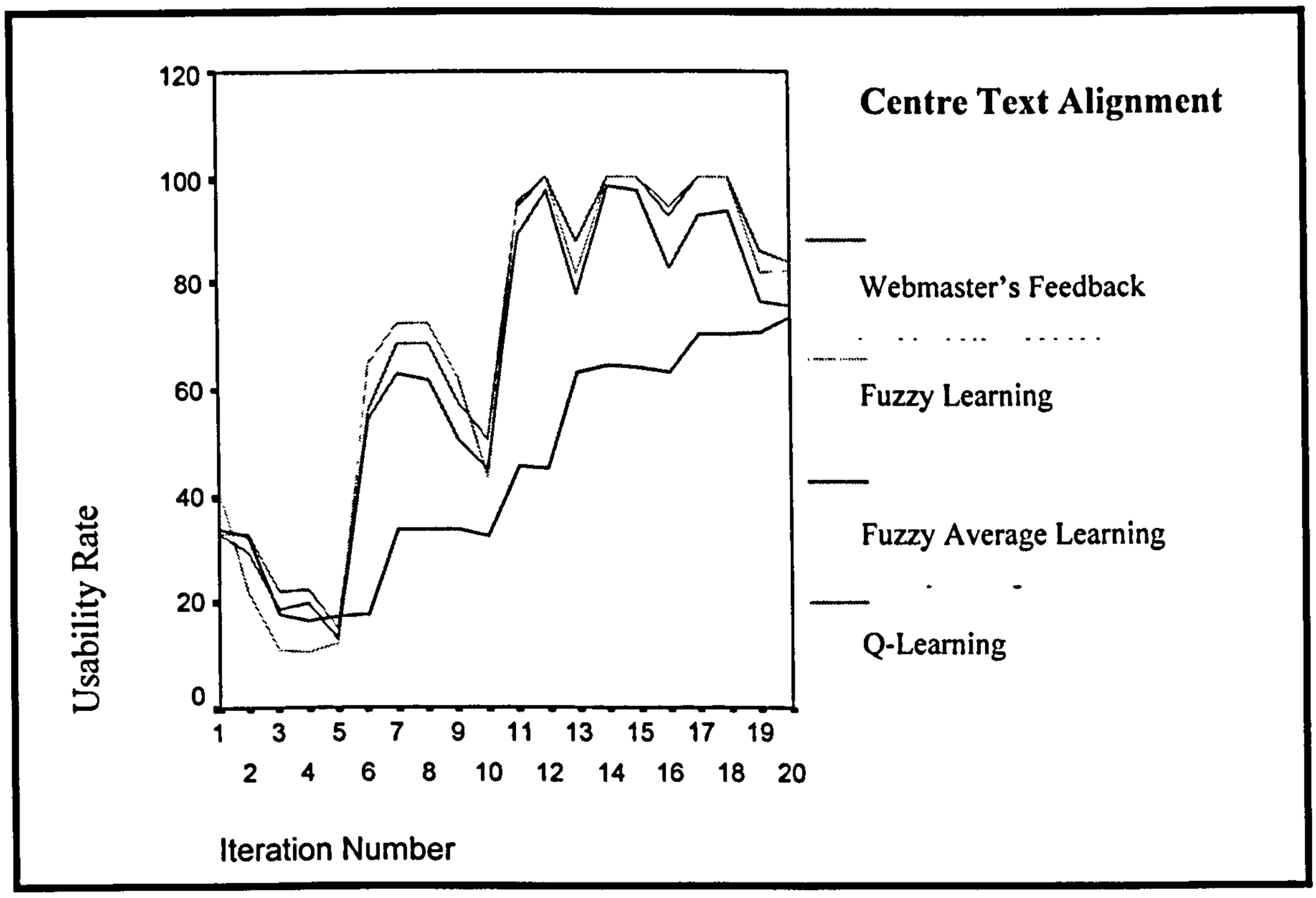


Figure B.11: The Customisation of the Centre Text Alignment Usability Rate in the BBC Web Page

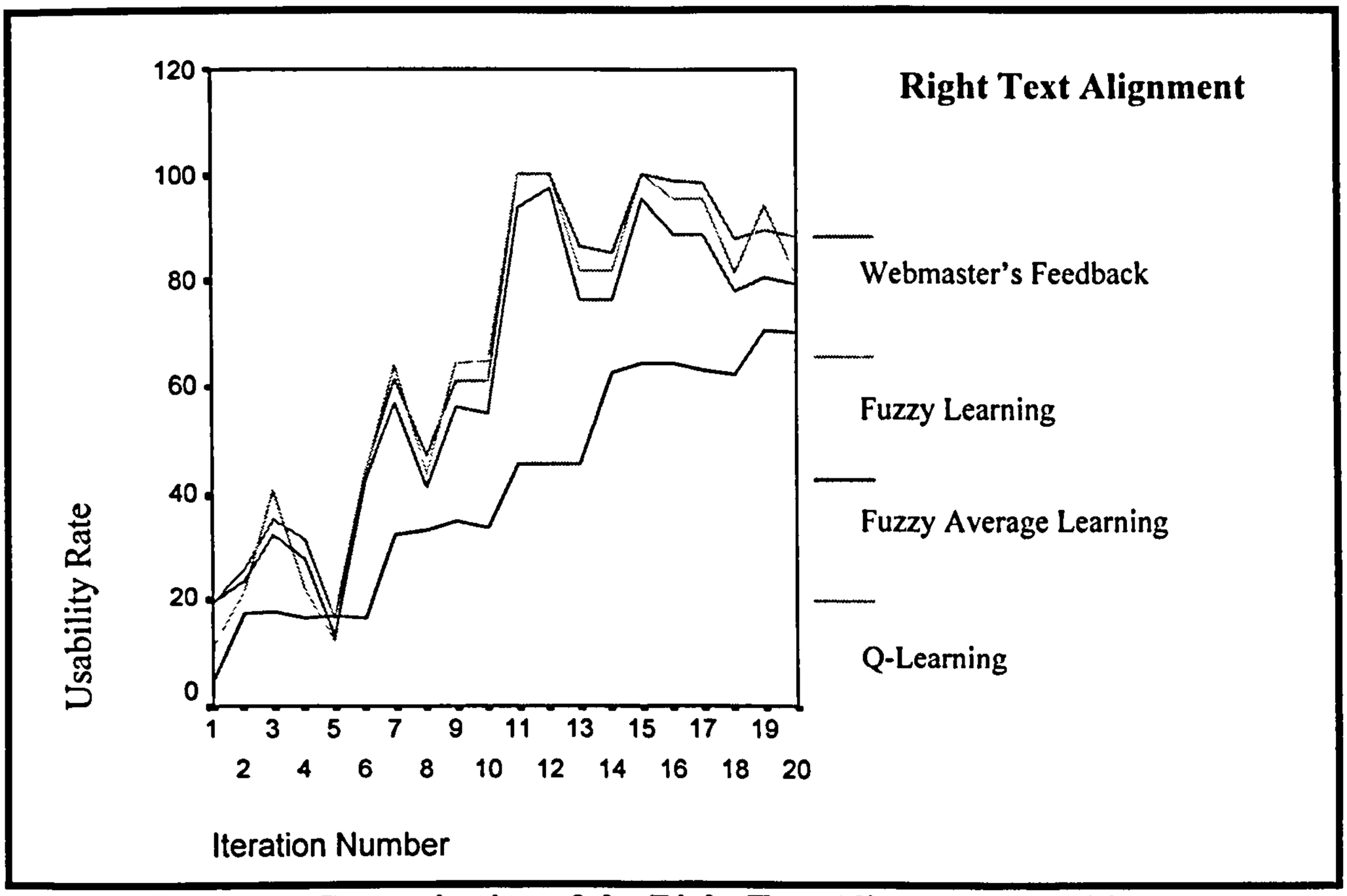


Figure B.12: The Customisation of the Right Text Alignment Usability Rate in the BBC Web Page

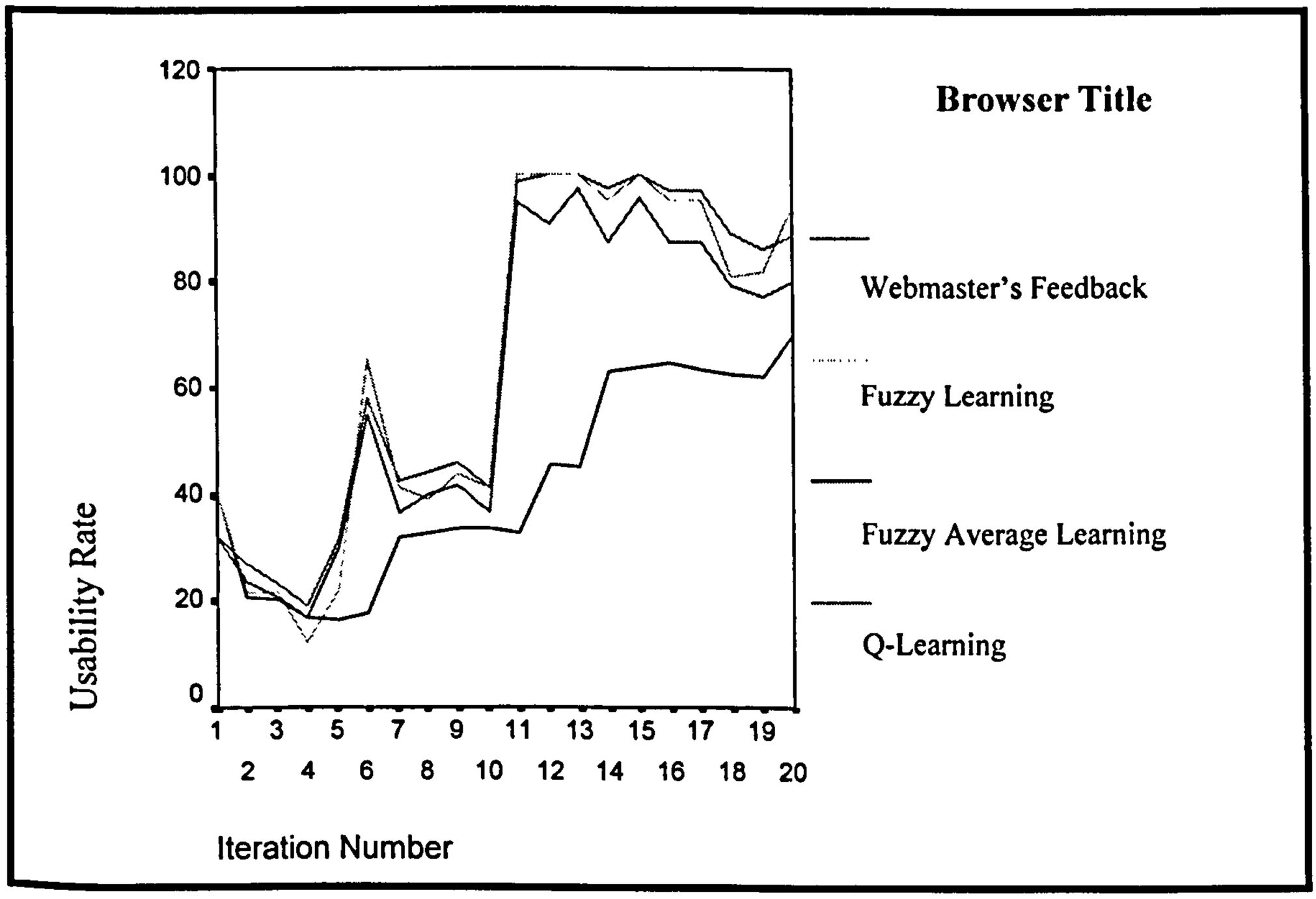


Figure B.13: The Customisation of the Browser Title Usability Rate in the BBC Web Page

The following tables represent the usability evaluation values produced by each learning approach for both Aljazeera and BBC web pages. The values presented during the 20 iterations of usability evaluation performed for each web page. The changes in the usability evaluation are customised to the webmaster's feedbacks presented in Table 9.1 and Table 9.2 illustrated in Chapter 9.

Appendix C

A Fuzzy Model to Measure Colour Contrast as an aspect of Web Usability

Maysoon Abulkhair

Department of Computer Science
University of Sheffield
Regent Court, 211 Portobello Street,
Sheffield S1 4DP
m.abulkhair@dcs.shef.ac.uk

Siobhán North

Department of Computer Science
University of Sheffield
Regent Court, 211 Portobello Street,
Sheffield S1 4DP
s.north@dcs.shef.ac.uk

Abstract

Colour is not only an aesthetic feature of web design; it can be employed to emphasize specific items and to send signals to the user. There are many aspects of effective colour use; the most obvious are the consistency of the colours, their number and the relationship between the text and background colours. This last is often the most important but all of them affect web usability to some extent.

This paper describes a fuzzy colouring model to measure objectively the relationship between the text and background and to produce a colour usability rating based on both web experts recommendation, to have the maximum contrast between the text and the background colours, and experimental results. This model is built in two phases, the intensity measuring phase and the colouring usability rate phase. The fuzzy model described here gives better usability rates than the colouring difference model suggested by other authors.

1 Introduction

Many web usability recommendations exist to guide web authors in the use of the colour. Most of these guidelines are based on the contrast between the text and the background colours, however almost all of them use very imprecise measures like "sufficient contrast" (Kerr, 2001; Sklar, 2000; Badre, 2002; Lynch and Horton, 2001), "high contrast" (Nielsen, 2000; Nielsen and Tahir, 2002; Rigden, 1999) etc.. These guidelines are inappropriate for automated usability measuring and there are too many pages on the web now to evaluate them consistently without some sort of automated tool. A more useful mechanism for measuring colour contrast in an appropriate way needs to be evolved. We have developed a scheme derived from the common web usability guidelines on colour but one which is quantifiable and testable.

There is one existing quantifiable procedure, suggested by (Manley, 2001), which consists of adding the three values of the red, green and blue (RGB) colour components together for both the text and the background and then finding the difference. According to Manley the colour difference should be at least 255 (decimal) to provide sufficient contrast between the text and the

background colour. This colour difference measure does provide a starting point but is too crude to be useful.

Our model involves measuring colour usability using fuzzy logic to incorporate the usual guidelines. It is constructed in two phases; the colour intensity phase and the colour usability rate phase. Both are described below in sections 2 and 3 and the results are discussed in sections 4 and 5.

2 Colour Intensity Phase

It is necessary to determine the colour intensity of both the text and the background colours prior to the colour contrast assessment, which in turn is necessary to measure the usability rate. This phase uses the fuzzy rules to infer the colour intensity from its RGB components and their variations. Each of these components can be split into five fuzzy sets: Dark, MidDark, Medium, MidLight and Light, according to the amount used by the web designer and then combined to give an overall colour intensity measure. To determine the overall colour intensity various fuzzy rules are applied to combine the RGB components in order to give a specific colour intensity value. The colour intensity is quantified as: Darkest, VeryDark, Dark, MildDark, Darkish, Medium, Lightish, MildLight, Light, VeryLight and Lightest. The fuzzy rules involve combining the RGB intensities determined earlier in the fuzzy sets to produce a fuzzy colour intensity. This process can be illustrated by the following example:

IF (Red IS Dark) AND (Green IS Dark) AND (Blue IS Dark) THEN
(ColourIntensity IS Darkest) **(1)**

Having established a general rule for colour intensity measuring, we also considered the special case of red/green colour blindness which is fairly common. The red/green colour blindness case is dealt with using the RGB colour intensities to produce an intensity of either Red or Green. This fuzzy rule is of the form:

IF (Red IS Dark) AND (Green IS Medium) AND (Blue IS Dark) THEN
(ColourIntensity IS Dark) AND (ColourBlind IS Green) **(2)**

There are around 125 fuzzy rules to derive the colour intensity measure, and all of them followed the form of either example (1) or (2). Both forms of fuzzy rule are based on experimental results. The colour intensity crisp value, computed from the deffuzzification process, varies from (0-100) where 0 is the darkest and 100 is the lightest.

3 Colour Usability Rate Phase

Once the colour intensities have been established they are employed in the colour usability rate phase. This is a crucial part of the model, because it integrates the two colours intensities with the most common colour blindness case to produce an appropriate colour usability rate. Thus the specification of the colour intensity fuzzy set is inherited from the previous phase to infer the colour usability rate from the colour contrast measure. The colour usability rate can be categorised as VeryHigh, High, MildHigh, Highish, Medium, Lowish, MildLow, Low and VeryLow. The VeryHigh colour usability rate is achieved by the most widely recommended usability guidelines for colour. Most of the fuzzy rules that determine the colour usability rate are derived from previous studies together with experimental results. For instance, the combination of white background colour together with black text (but not the other way around) will be given the highest colour usability rate (Nielsen, 2000; Nielsen & Tahir, 2002; Lynch & Horton, 2001; Shneiderman, 1998) whereas a low contrast between the text and background colour indicates low colour usability evaluation. This guideline is applied in the following form of the fuzzy rules:

IF (TextIntensity IS Darkest) AND (BackGroundIntensity IS Lightest) THEN
(ColourUsabilityRate IS VeryHigh)

Even where there is sufficient contrast between a dark background and light text, some usability experts give this a lower usability rate than the reverse (Nielsen, 2000; Nielsen & Tahir, 2002; Lynch & Horton, 2001; Shneiderman, 1998). However, there are other usability guidelines which strongly recommend the use of the dark background, light text combination (Preece et al 1994; Rivlin, Lewis & Davies-Cooper 1990). After some experiment, it was decided to adopt the more recent recommendations and the colour contrast fuzzy model will give this a lower usability rate than the dark background/light text combination. The corresponding fuzzy rule used to assess this is in the form:

IF (TextIntensity IS Lightest) AND (BackGroundIntensity IS Darkest) THEN (ColourUsabilityRate IS High)

Whereas, low contrast between the text and background colour results in a low colour usability evaluation. This gives us a group of rules as follows:

IF (TextIntensity IS Dark) AND (BackGroundIntensity IS Dark) THEN (ColourUsabilityRate IS VeryLow)

IF (TextIntensity IS Light) AND (BackGroundIntensity IS Lightest) THEN (ColourUsabilityRate IS VeryLow)

IF (TextIntensity IS Medium) AND (BackGroundIntensity IS Darkish) THEN (ColourUsabilityRate IS VeryLow)

There are more than 140 fuzzy rules in the model to cover different aspects of colour contrast. These fuzzy rules are associated with each other to give a reliable colour usability rate. The crisp value, produced from the deffuzzification process, again varies from (0 – 100) where 0 is the lowest and 100 is the highest colour usability rate.

4 Evaluation

The fuzzy model has been used to produce a reliable colour usability rate. For example, the combination (white text on green background) gives a usability rate of 12.2% (Low) but the calculated intensity of the background is 77 (Light) and the text intensity is 100 (Lightest). So, in the colour difference model, it would be acceptable at 33.3%. The other example, involving the most common form of colour blindness, with the combination of green shade (# 99 ff 99) text and red (# ff 00 00) as a background is acceptable in the colour difference model and rated at 40% whereas, the fuzzy model produces 23.7% colour usability rate.

Unlike the previous models for calculating the contrast between the text and the background colours, this fuzzy model deals with the combination differently. It is not necessary for the same combination to have the same usability rate when exchanging the text with the background colour and vice versa. As discussed earlier, black text on a white background is not the same as white text over a black background, so, the highest usability rate, 95%, is given to the black text over white colour combination by our fuzzy model whereas reversing this combination to black text on a white background is only rated as 83.3%.

There are some other combinations that commonly annoy users even with normal vision, and, as result they might face difficulties when reading them. For example, light pink text (# ffccff) over green background (#00ff00) where both colours picked from the safe web colours, is given a rating of 34%, but even though it provides a reasonable degree of contrast. With the difference model mentioned earlier this would be given 60%.

5 Conclusion and Future Work

Our fuzzy model considers the most recent and common usability recommendations and transforms them into a consistent and quantifiable form suitable for automated evaluation of web pages. This is, in itself, useful but an even more useful feature of our model is that it can easily be

adapted to take into account other aspects of colour as it affects usability. We intend to extend the model to include further experimental results on different users' and cultural groups' preferences in terms of colour. It can of course also be adapted to take into account other researchers results in the same area as they emerge in the future.

References

Badre A. (2002). Shaping Web Usability Interaction Design in Context. Boston: Addison Wesley.

Kerr, M. (2001). Tips and Tricks for Website Mangers. London: Aslib-IMI.

Lynch, P., Horton, S. (2001). Web Style Guide: Basic Design Principles for Creating Web Sites (2nd ed.) New Haven: Yale University Press.

Manley B. (2001), quoted in Tips and Tricks for Website Mangers. Edited by Mark Kerr. London: Aslib-IMI.

Nielsen J. (2000). Designing Web Usability: The Practice of Simplicity, Indianapolis: New Rider publishing.

Preece J., Rogers Y., Sharp H., Benyon D., Holland S., Carey T. (1994). Human-Computer Interaction. Harlow, England: Addison-Wesley.

Rigden, C. (1999). 'The Eye of the Beholder'- Designing for Colour-Blind Users. *British Telecommunications Engineering*, Vol.17.

Rivlin, C., Lewis R., Davies-Cooper R. (1990). Guidelines for Screen Design. London: Blackwell Scientific publications.

Shneiderman B. (1998). Designing the User Interface Strategics for Effective Human-Computer Interaction (3rd ed). Reading, MA: Addison-Wesley.

Sklar, J. (2000). Principles of Web Design. Cambridge, MA: Course Technology.

Tahir M., Nielsen J. (2002). Homepage Usability 50 Websites Deconstructed. Salem, Virginia: New Rider publishing.

Appendix D

User Satisfaction Questionnaires

D.1 Colour Usability Questionnaire (English version)

Have you realised that the colours used can either draw your eyes towards or repel them from a specific web page? So how much do colours matter in web page design? This questionnaire is part of an attempt to find the answer. Please help me by filling it in.

Question 1

What is your native language?

- ○ Arabic
- ○ English
- ○ Other

Question 2

Which of the following is the colour combination you like?

- ○ Dark-Blue text and White background
- ○ White text and Dark-Blue background
- ○ Black text and Light-Yellow background
- ○ Yellow text and Black background

Question 3

Do you like subtitles and the important information in the text to be in other colours?

- ○ Yes
- ○ No

Question 4

What are the areas that you prefer to be in different colours?

- ○ Document Title
- ○ Important information in the text
- ○ Tables
- ○ Subtitles

○ None

Question 5

Do you like a Dark background colour?

○ Yes

○ No

Question 6

Which is closest word to what the following colours suggest to you?

White [Purity]

Black [Mourning]

Red [Attention]

Blue [Sad]

Yellow [Happiness]

Green [Envy]

Question 7

What do the following combinations remind you of?

Red text and **White** background [Joy]

Blue text and **White** background [Nationalism]

Green text and **White** background [Peace]

Yellow text and **Black** background [Bugs]

Green text and **Red** background [Christmas]

In the following questions the same 4 examples will be used

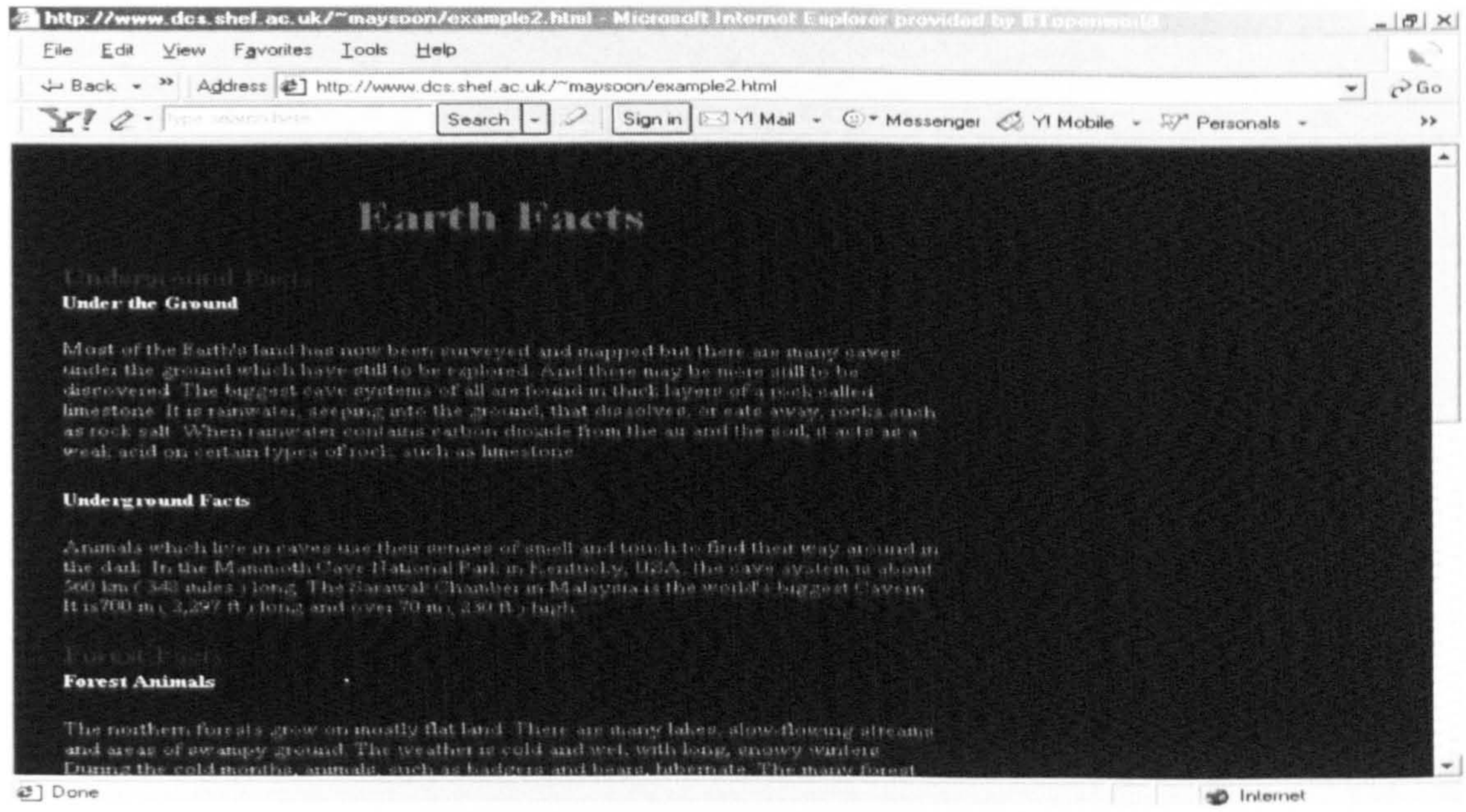
Question 8

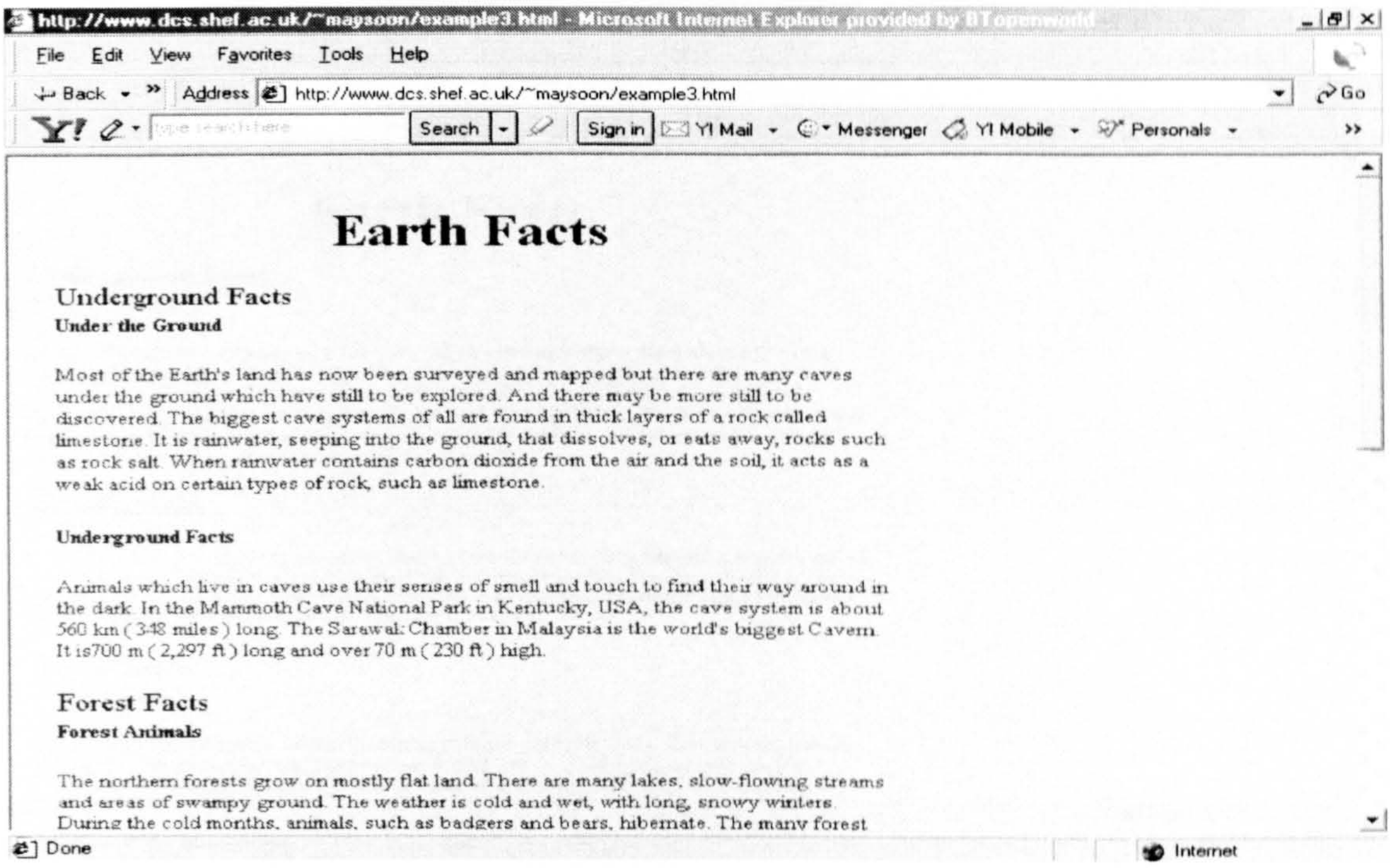
Which of these 4 samples do you prefer? Arrange them in order marking your favourite as first and so on

First

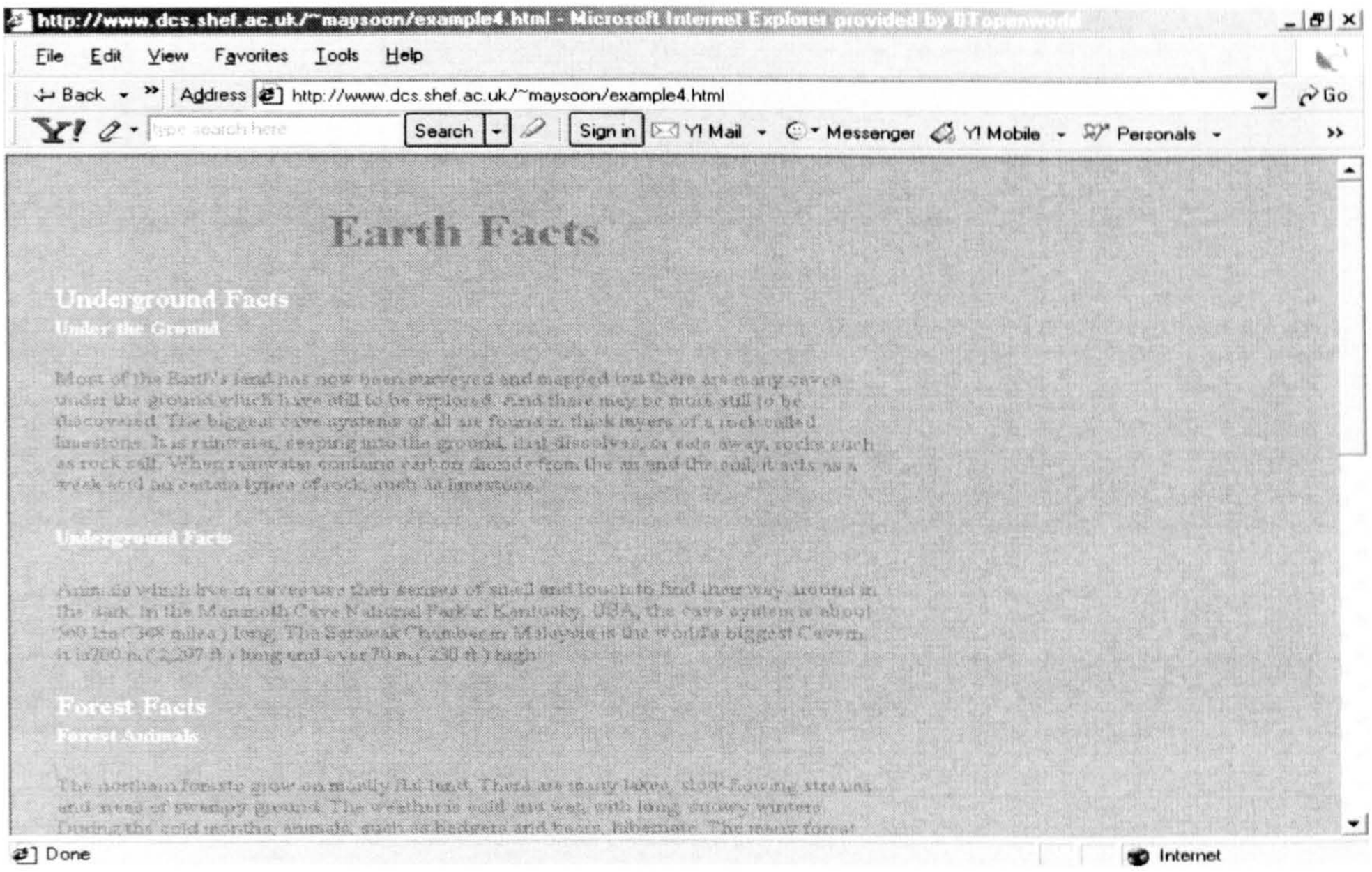

http://www.dcs.shef.ac.uk/~maysoon/example1.html - Microsoft Internet Explorer provided by
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/example1.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
Earth Facts
Underground Facts
Under the Ground
Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.
Underground Facts
Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.
Forest Facts
Forest Animals
The northern forests grow on mostly flat land. There are many lakes, slow-flowing streams and areas of swampy ground. The weather is cold and wet, with long, snowy winters. During the cold months, animals, such as badgers and bears, hibernate. The many forest
Done
Internet

First


http://www.dcs.shef.ac.uk/~maysoon/example2.html - Microsoft Internet Explorer provided by
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/example2.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
Earth Facts
Under the Ground
Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.
Underground Facts
Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.
Forest Facts
Forest Animals
The northern forests grow on mostly flat land. There are many lakes, slow-flowing streams and areas of swampy ground. The weather is cold and wet, with long, snowy winters. During the cold months, animals, such as badgers and bears, hibernate. The many forest
Done
Internet



First



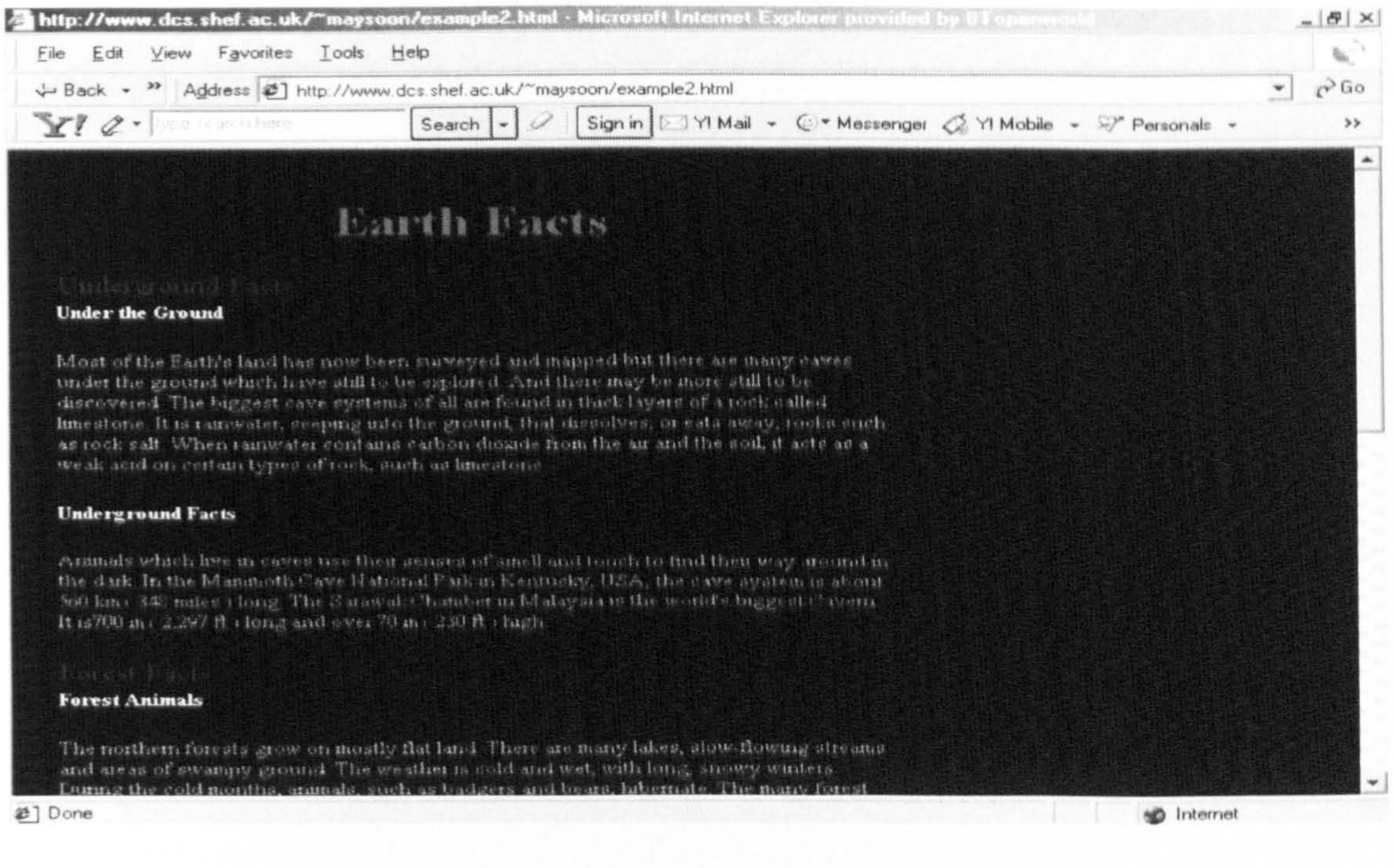
Question 9

Please rate the attractivness of the colours selected in the following designs as

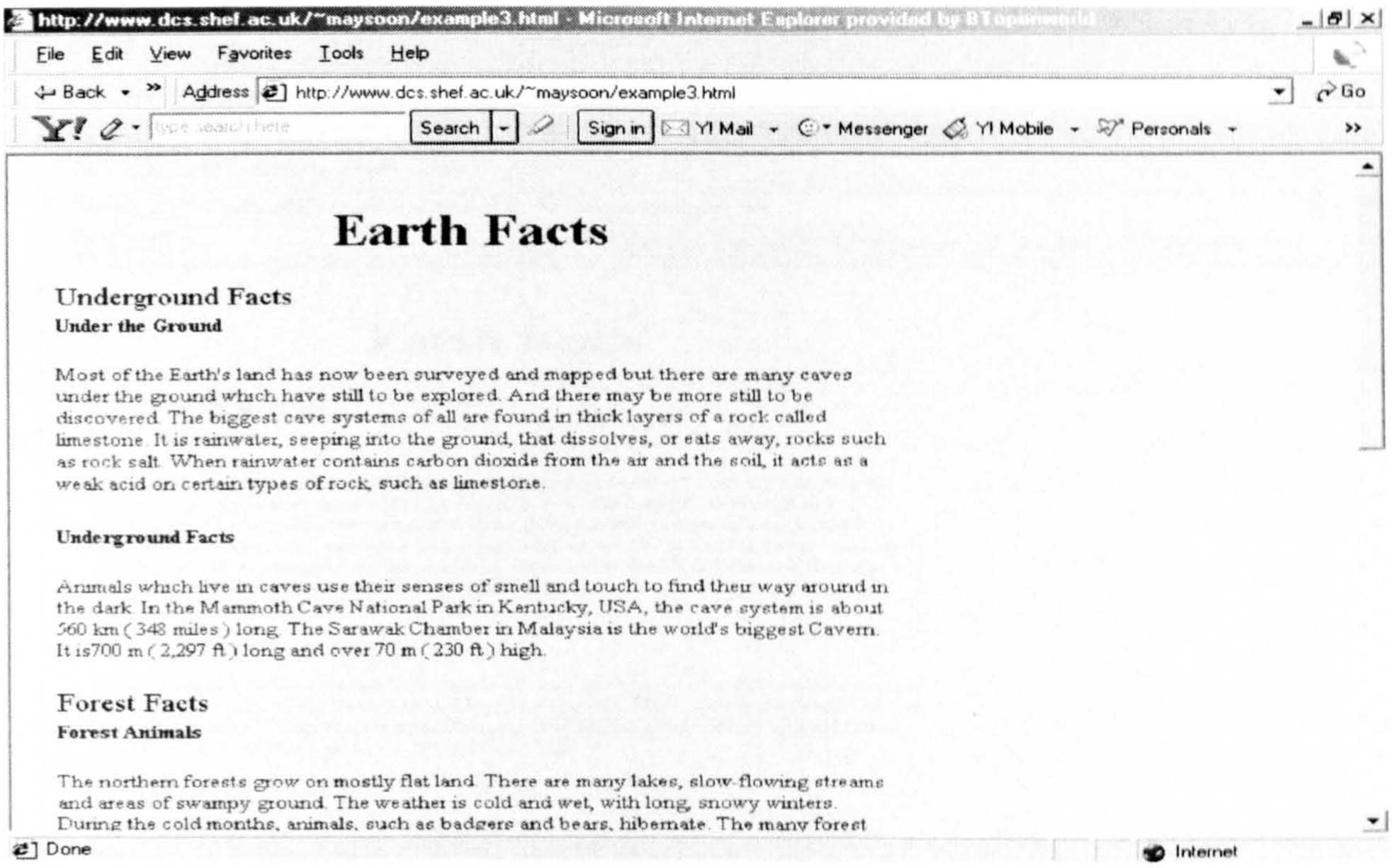
Good


http://www.dcs.shef.ac.uk/~maysoon/example1.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/example1.html
Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
Earth Facts
Underground Facts
Under the Ground
Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.
Underground Facts
Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.
Forest Facts
Forest Animals
The northern forests grow on mostly flat land. There are many lakes, slow-flowing streams and areas of swampy ground. The weather is cold and wet, with long, snowy winters. During the cold months, animals, such as badgers and bears, hibernate. The many forest
Done
Internet

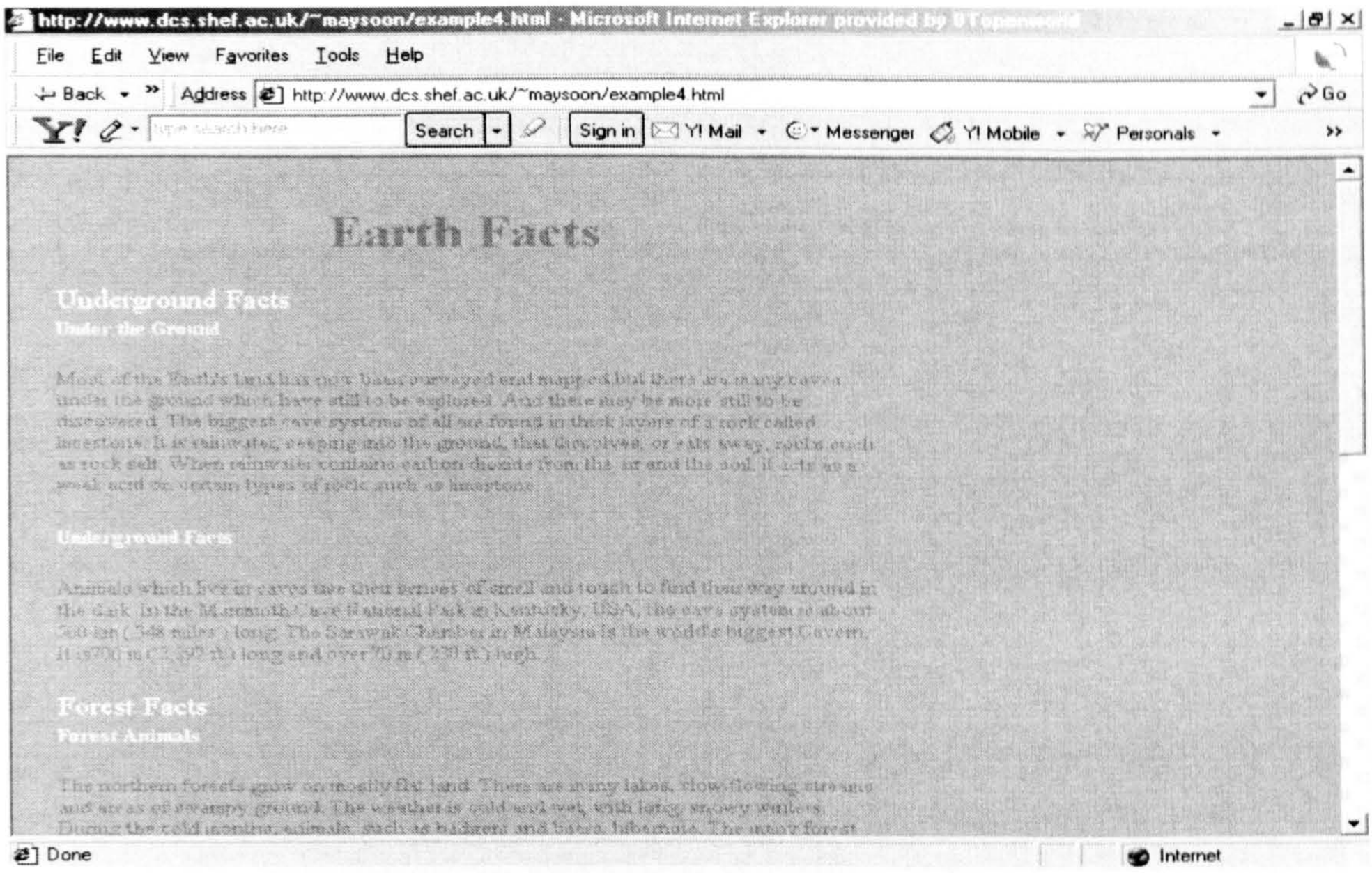
Good


http://www.dcs.shef.ac.uk/~maysoon/example2.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/example2.html
Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
Earth Facts
Under the Ground
Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.
Underground Facts
Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.
Forest Facts
Forest Animals
The northern forests grow on mostly flat land. There are many lakes, slow-flowing streams and areas of swampy ground. The weather is cold and wet, with long, snowy winters. During the cold months, animals, such as badgers and bears, hibernate. The many forest
Done
Internet

Good



Good



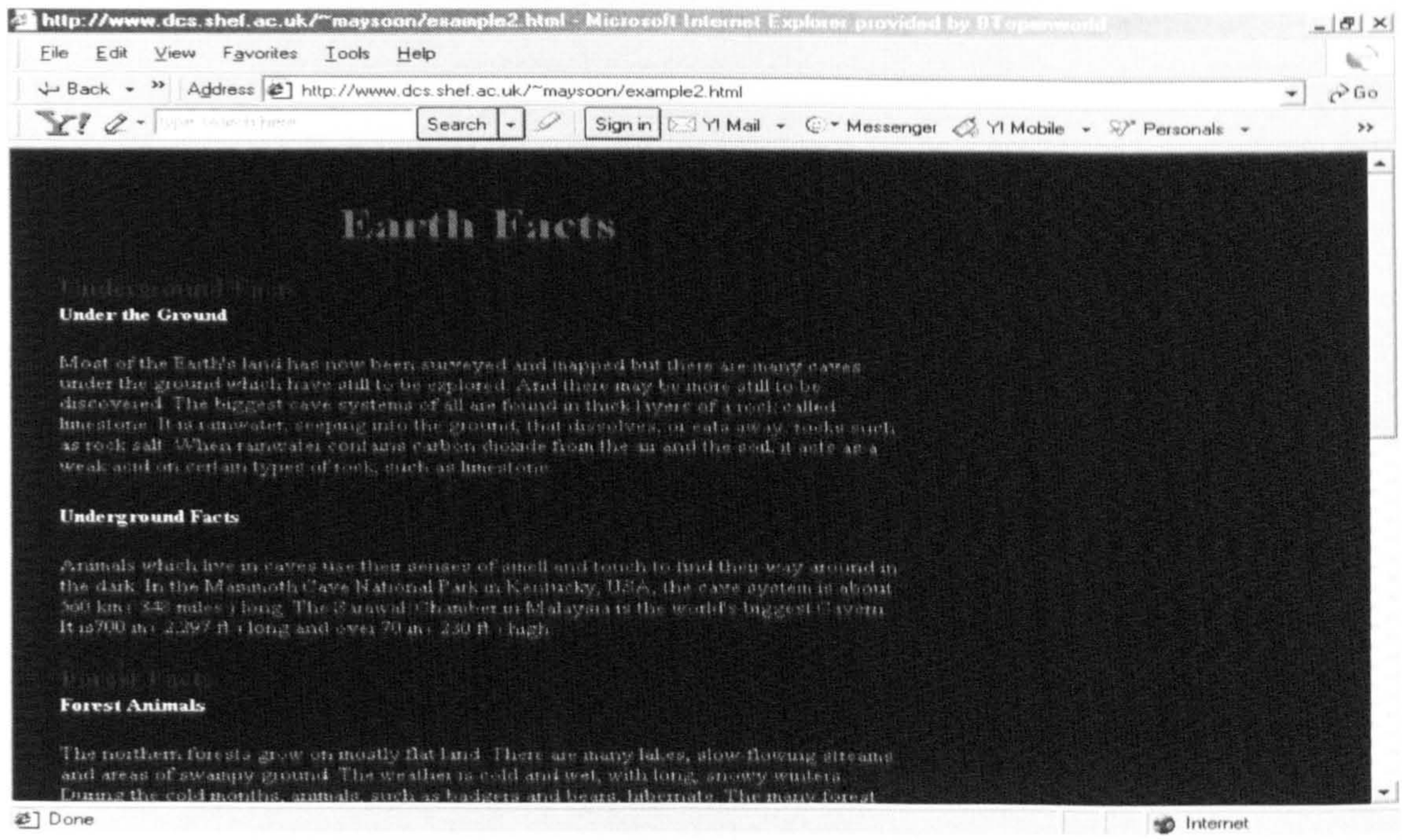
Question 10

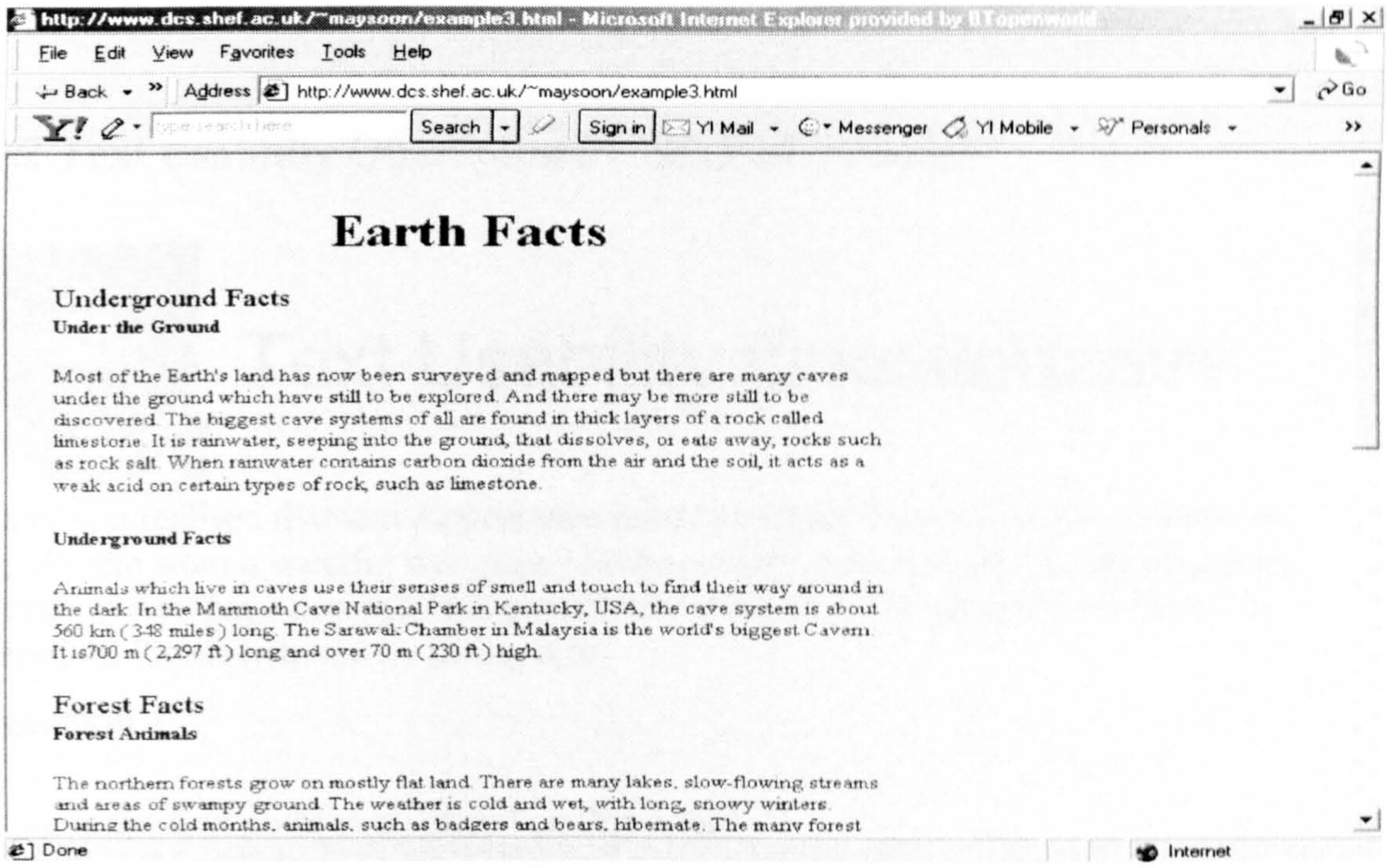
Please rate the readability of the colours selected in the following designs

Good

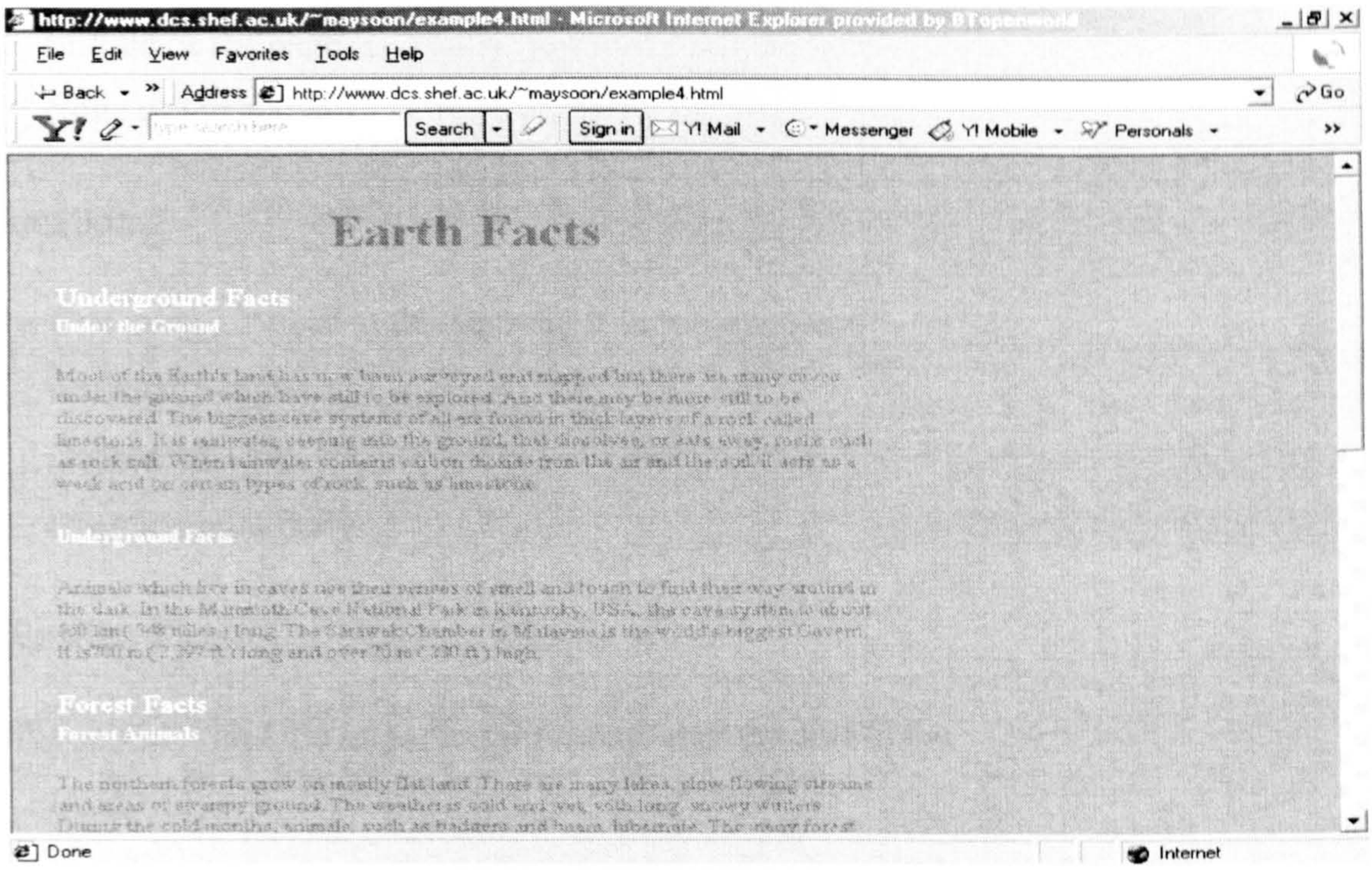

http://www.dcs.shef.ac.uk/~maysoon/example1.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/example1.html
Go
Search
Sign in
Y! Mail
Messenger
Y! Mobile
Personals
Earth Facts
Underground Facts
Under the Ground
Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.
Underground Facts
Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.
Forest Facts
Forest Animals
The northern forests grow on mostly flat land. There are many lakes, slow-flowing streams and areas of swampy ground. The weather is cold and wet, with long, snowy winters. During the cold months, animals, such as badgers and bears, hibernate. The many forest
Done
Internet

Good


http://www.dcs.shef.ac.uk/~maysoon/example2.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/example2.html
Go
Search
Sign in
Y! Mail
Messenger
Y! Mobile
Personals
Earth Facts
Under the Ground
Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.
Underground Facts
Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.
Forest Facts
Forest Animals
The northern forests grow on mostly flat land. There are many lakes, slow-flowing streams and areas of swampy ground. The weather is cold and wet, with long, snowy winters. During the cold months, animals, such as badgers and bears, hibernate. The many forest
Done
Internet



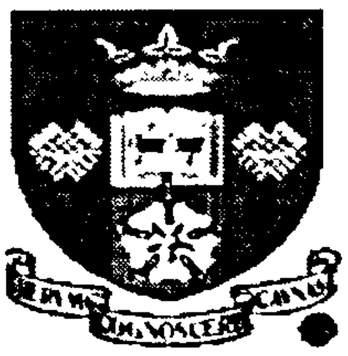
Good



reset

Thank you for your time and the valuable information you are submitting.

D.2 Text Usability Questionnaire (English version)



Text Usability Questionnaire

Have you realised that text Appearance used can either draw your eyes towards or repel them from a specific web page? So how much does text and its specifications matter in web page design? This questionnaire is part of an attempt to answer the question. Please help me by filling it in.

Question 1

What is the first point you noticed on this page?

- ○ Upper left corner
- ○ Upper center
- ○ Upper right corner
- ○ Middle of the page
- ○ None of the above

Question 2

The first thing I looked at deliberately it in this page was:

- ○ University logo
- ○ Document Title
- ○ Questionnaire body

Question 3

I like subtitles and the important information in the text to be:

- ○ Bold
- ○ In italics
- ○ Underlined
- ○ A larger letter size
- ○ In capitals
- ○ In other colour

Question 4

Which of the following do you think can be scanned faster?

- Bold headings
- Italic headings
- Underlined headings
- Larger letter size headings
- Capital headings
- Coloured headings

Question 5

On basis of readability, decide how pleasant to read each of the 4 samples

Pleasant

Pleasant

http://www.dcs.shef.ac.uk/~maysoon/emphasizeexample2.html - Microsoft Internet Explorer provided by BT openworld

File Edit View Favorites Tools Help

Back Address http://www.dcs.shef.ac.uk/~maysoon/emphasizeexample2.html Go

Search Sign in Y! Mail Messenger Y! Mobile Personals

Earth Facts

Underground Facts

Under the Ground

Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.

Underground Facts

Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.

Last Updated: 6-09-2002 Important Notices

Done Internet

Pleasant

http://www.dcs.shef.ac.uk/~maysoon/emphasizeexample3.html - Microsoft Internet Explorer provided by BT openworld

File Edit View Favorites Tools Help

Back Address http://www.dcs.shef.ac.uk/~maysoon/emphasizeexample3.html Go

Search Sign in Y! Mail Messenger Y! Mobile Personals

Earth Facts

Underground Facts

Under the Ground

Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.

Underground Facts

Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.

Last Updated: 6-09-2002 Important Notices

Done Internet

Question 6

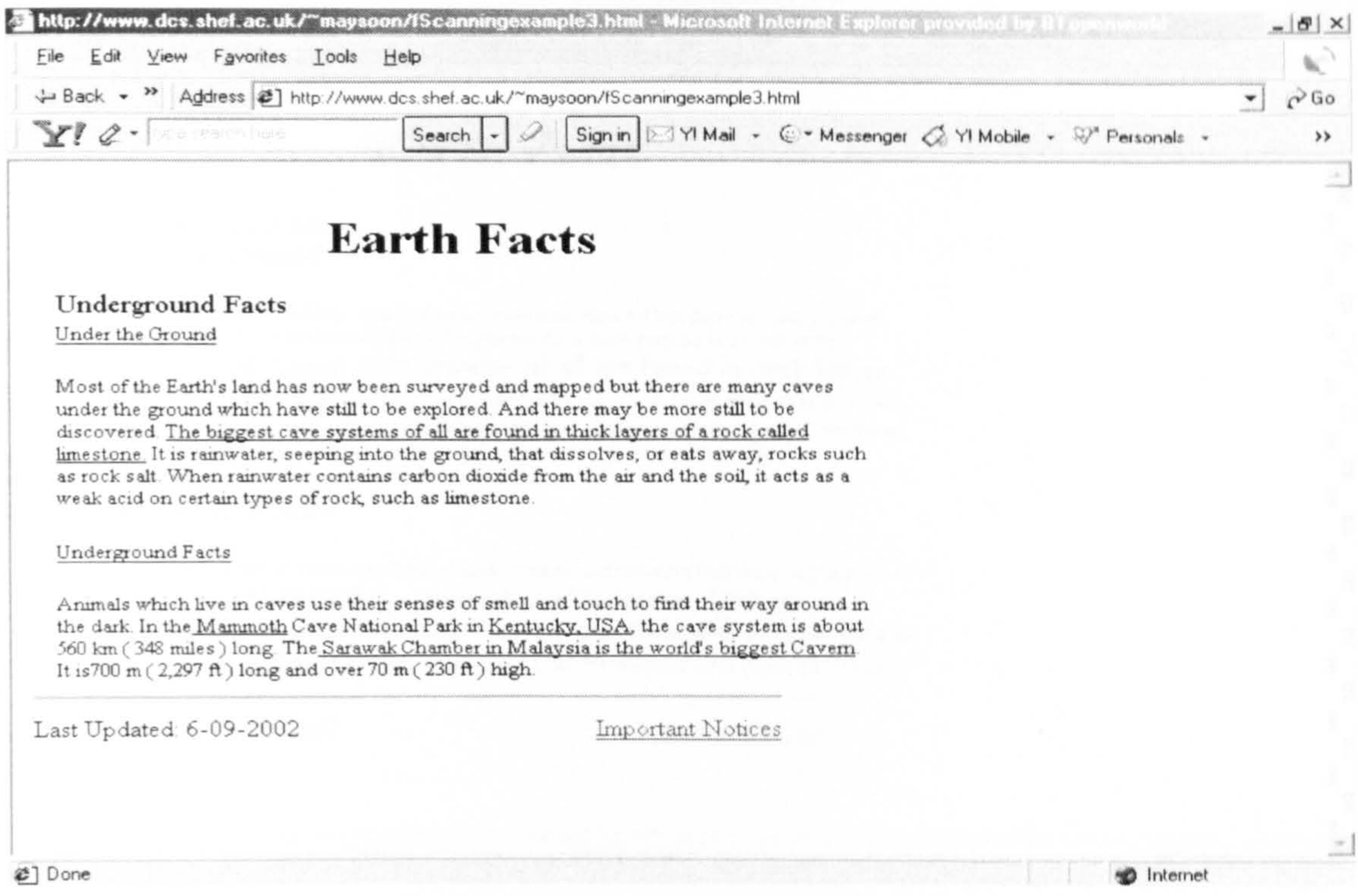
Determine how fast you can scan each of the following

Easy

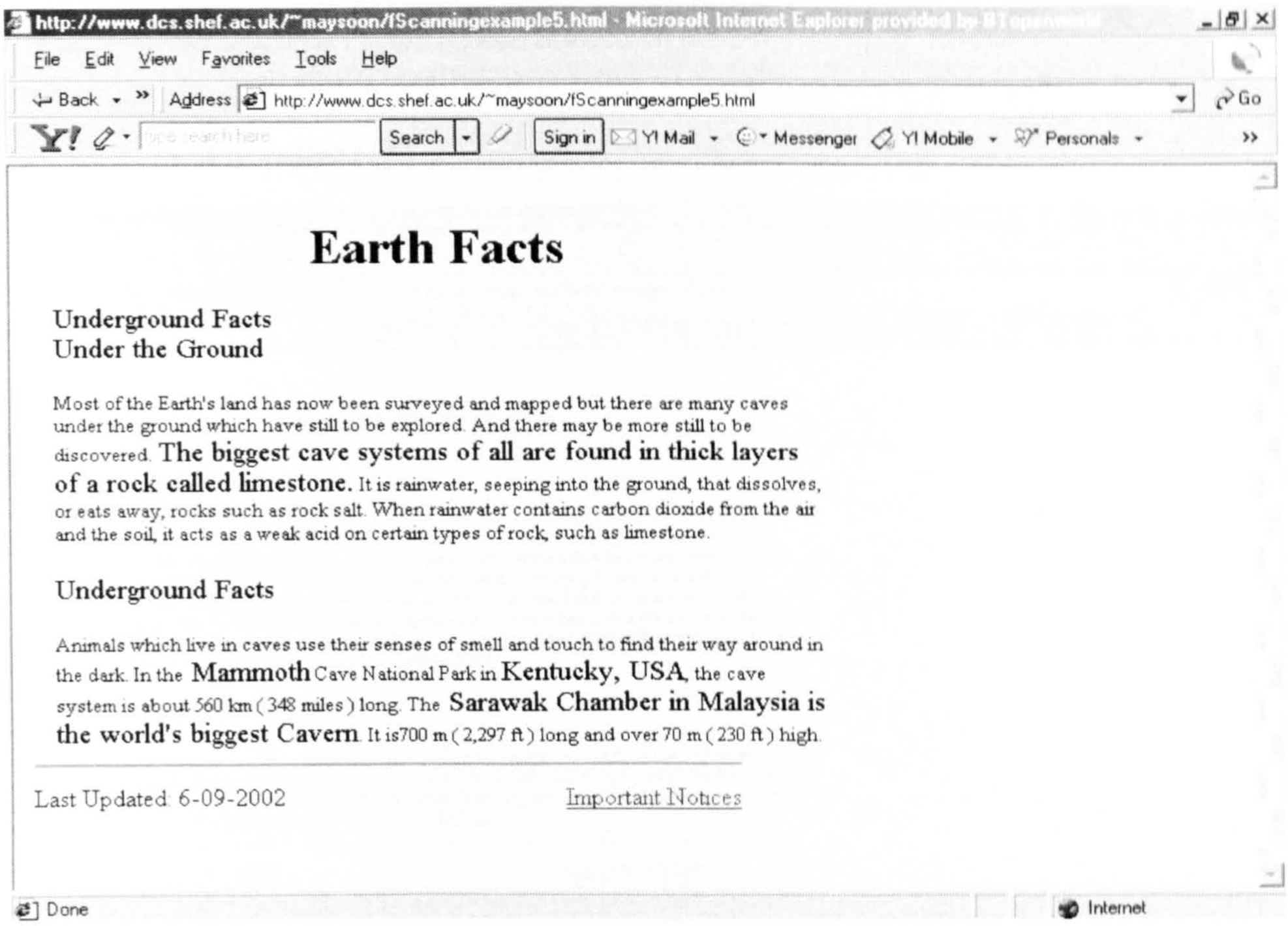
Easy


http://www.dcs.shef.ac.uk/~maysoon/fScanningexample2.html - Microsoft Internet Explorer provided by BT
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/fScanningexample2.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
Earth Facts
Underground Facts
Under the Ground
Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.
Underground Facts
Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.
Last Updated: 6-09-2002
Important Notices
Done
Internet

Easy


http://www.dcs.shef.ac.uk/~maysoon/fScanningexample3.html - Microsoft Internet Explorer provided by BT
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/fScanningexample3.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
Earth Facts
Underground Facts
Under the Ground
Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. The biggest cave systems of all are found in thick layers of a rock called limestone. It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.
Underground Facts
Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the Mammoth Cave National Park in Kentucky, USA, the cave system is about 560 km (348 miles) long. The Sarawak Chamber in Malaysia is the world's biggest Cavern. It is700 m (2,297 ft) long and over 70 m (230 ft) high.
Last Updated: 6-09-2002
Important Notices
Done
Internet

Easy

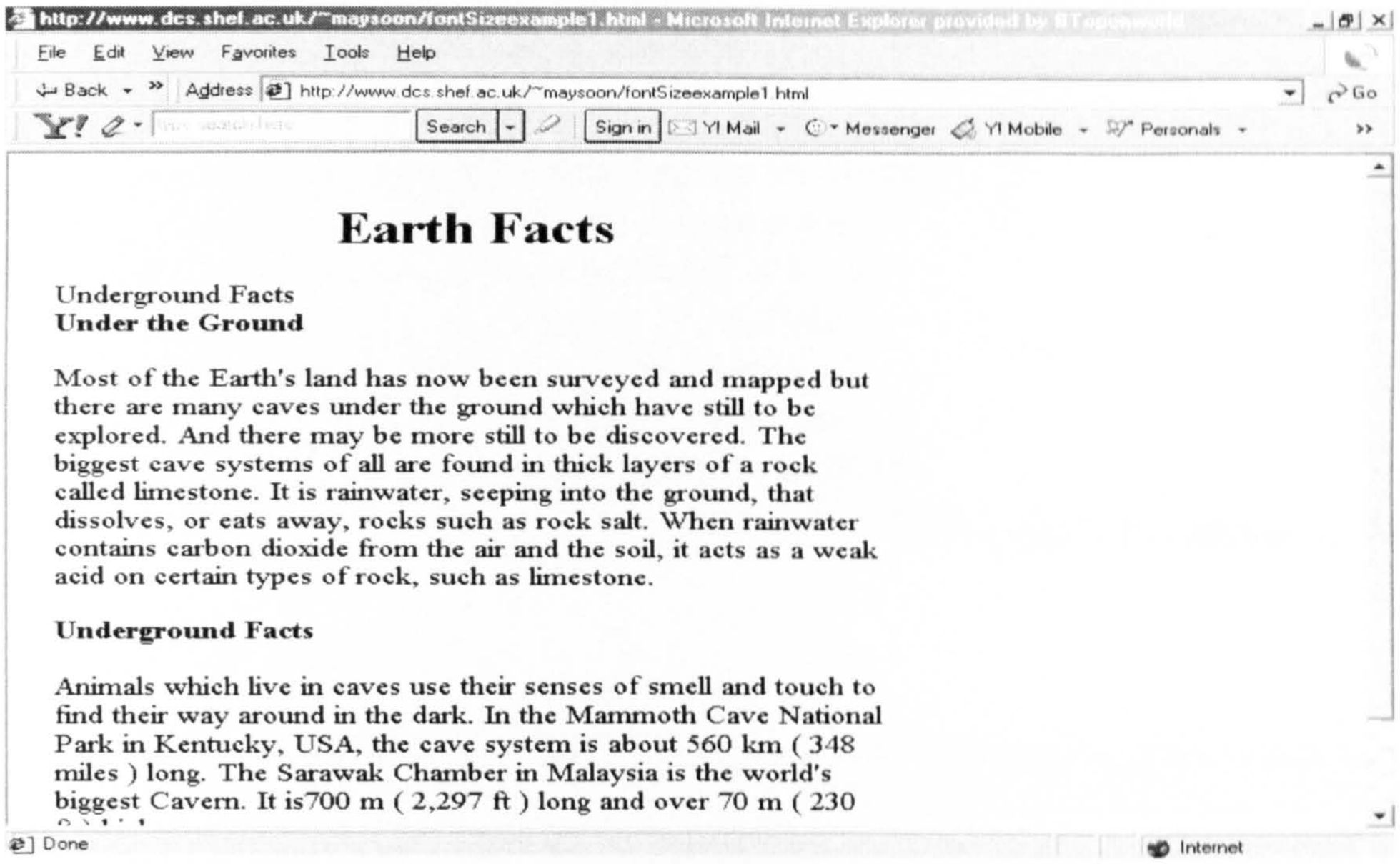


Question 7

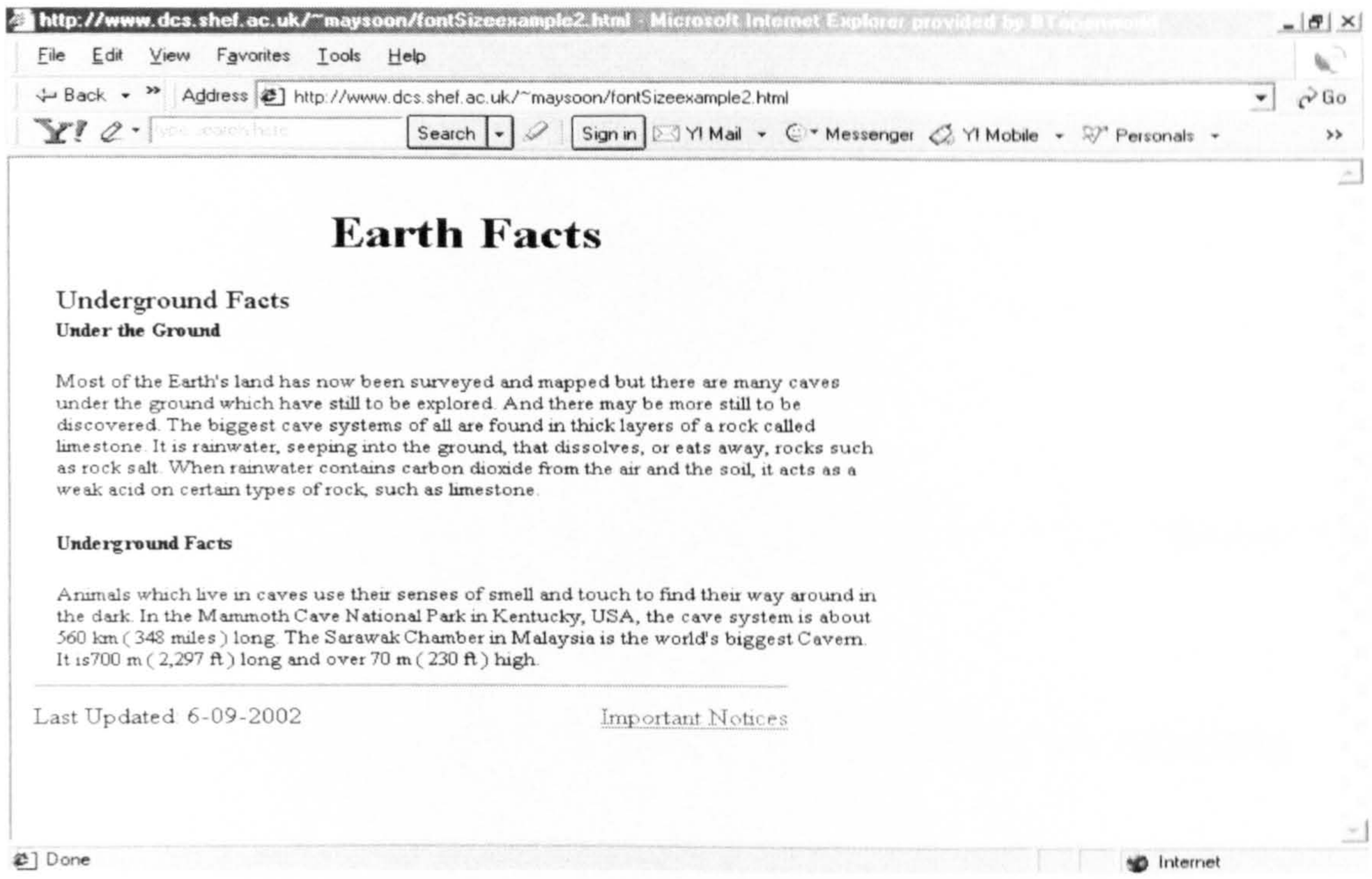
Question 7

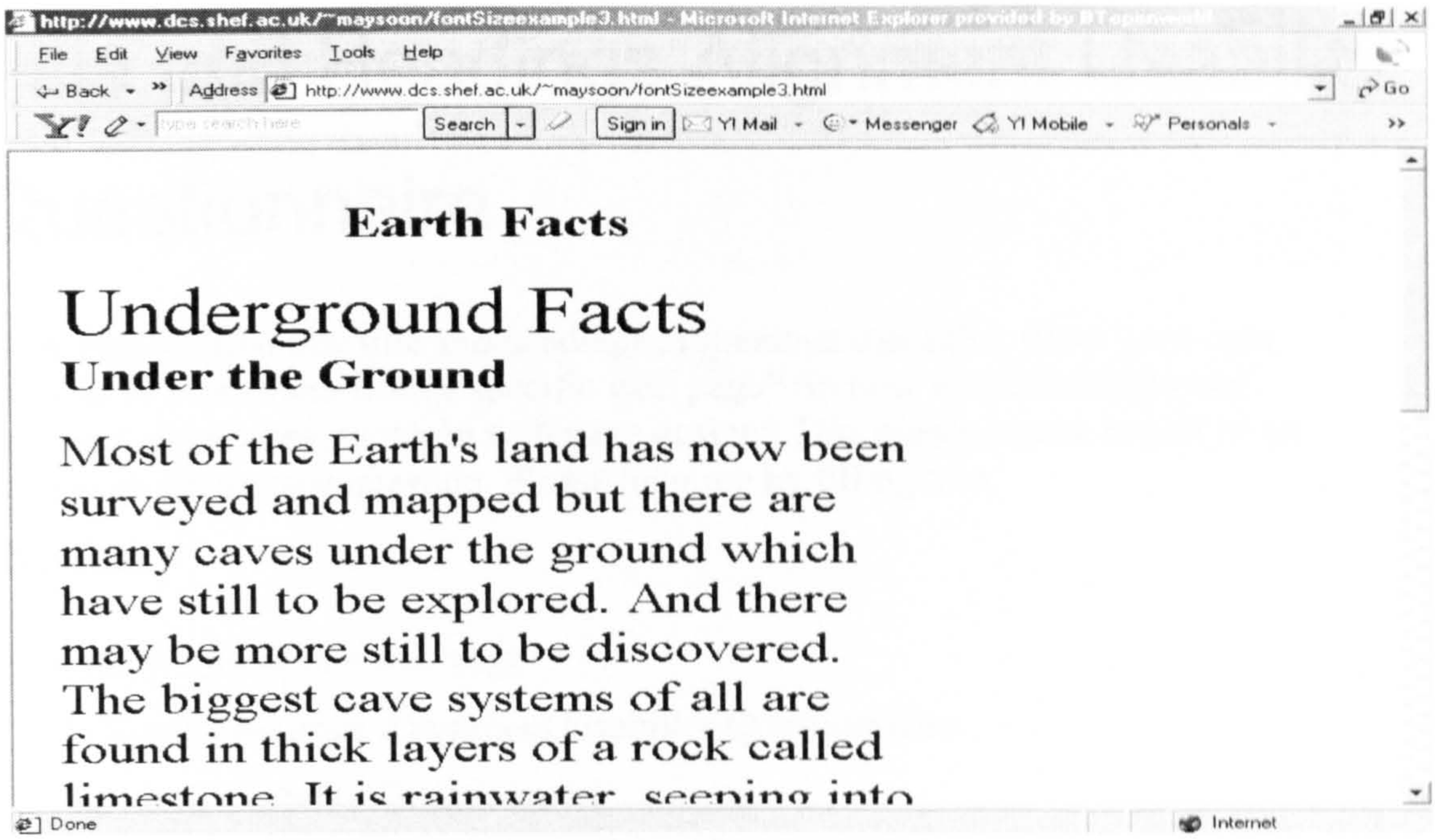
Determine the legibility level for each of the 4 samples

Legible



Legible





Legible

reset

Thank you for your time and the valuable information you are submitting.

D.3 Title Usability Questionnaire (English version)

Title and Headings Alignment Usability Questionnaire

Have you realised that title and headings alignments can either draw your eyes towards or repel them from a specific web page? So how much does title and headings alignments matter in web page design? This questionnaire is part of an attempt to answer the question. Please help me by filling it in.

Question 1

What is the title of this web page?

- Title and Headings Alignment Usability Questionnaire
- Research Usability Questionnaire

Question 2

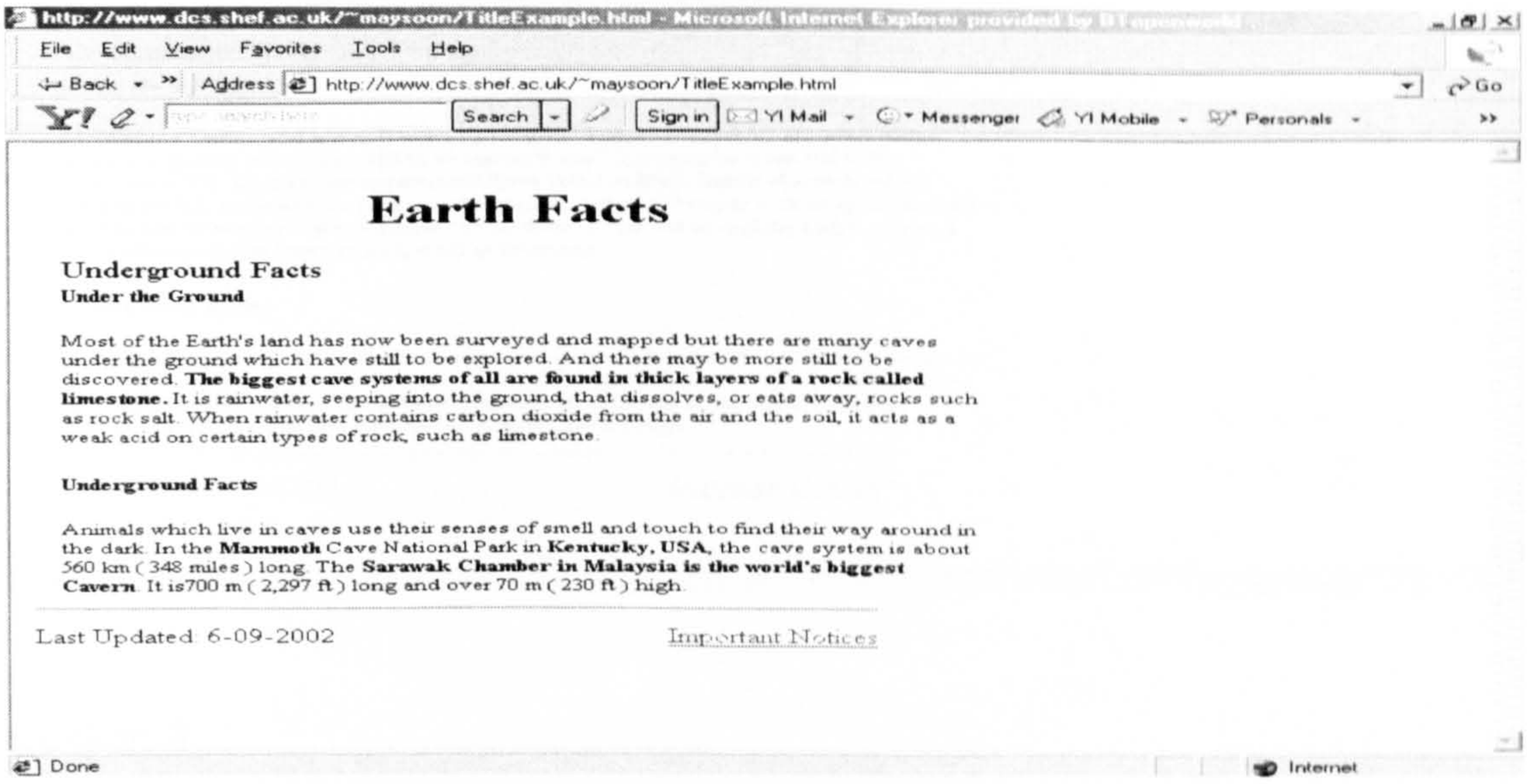
Do you like many links in a web page?

- Yes
- No

Question 3

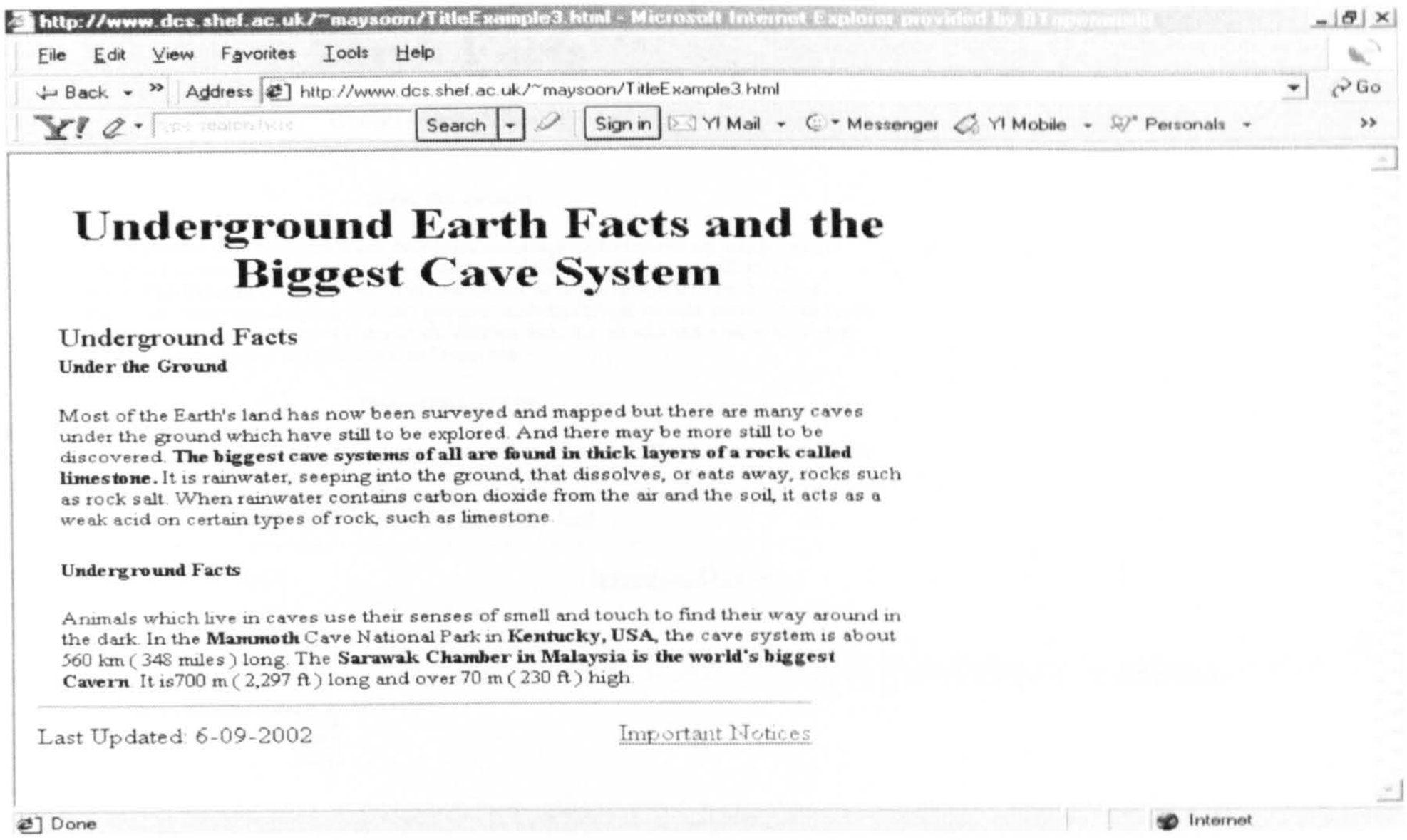
Based on memorability, evaluate each of the following titles

Memorable



Good

Legible



Question 4

Rate each of the 3 web pages based on how well organised they are

Organised

http://www.dcs.shef.ac.uk/~maysoon/AlignmentExample1.html - Microsoft Internet Explorer provided by BTopenworld

File Edit View Favorites Tools Help

Back Address http://www.dcs.shef.ac.uk/~maysoon/AlignmentExample1.html Go

Search Sign in Y! Mail Messenger Y! Mobile Personals

Earth Facts

Underground Facts

Under the Ground

Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. **The biggest cave systems of all are found in thick layers of a rock called limestone.** It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.

Underground Facts

Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the **Mammoth** Cave National Park in **Kentucky, USA**, the cave system is about 560 km (348 miles) long. The **Sarawak Chamber in Malaysia is the world's biggest Cavern**. It is700 m (2,297 ft) long and over 70 m (230 ft) high.

Last Updated: 6-09-2002 Important Notices

Done Internet

Organised

http://www.dcs.shef.ac.uk/~maysoon/AlignmentExample2.html - Microsoft Internet Explorer provided by BTopenworld

File Edit View Favorites Tools Help

Back Address http://www.dcs.shef.ac.uk/~maysoon/AlignmentExample2.html Go

Search Sign in Y! Mail Messenger Y! Mobile Personals

Earth Facts

Underground Facts

Under the Ground

Most of the Earth's land has now been surveyed and mapped but there are many caves under the ground which have still to be explored. And there may be more still to be discovered. **The biggest cave systems of all are found in thick layers of a rock called limestone.** It is rainwater, seeping into the ground, that dissolves, or eats away, rocks such as rock salt. When rainwater contains carbon dioxide from the air and the soil, it acts as a weak acid on certain types of rock, such as limestone.

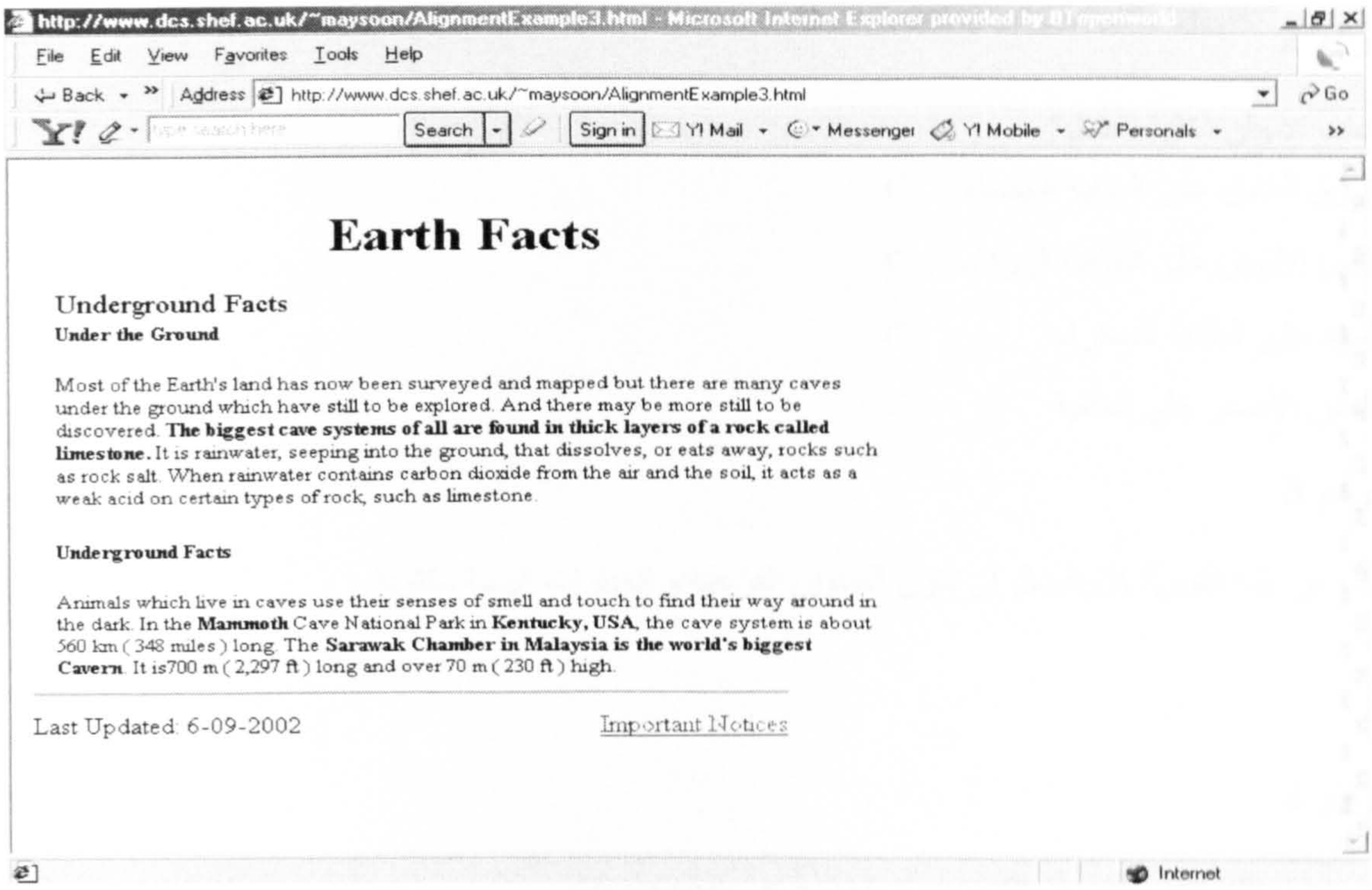
Underground Facts

Animals which live in caves use their senses of smell and touch to find their way around in the dark. In the **Mammoth** Cave National Park in **Kentucky, USA**, the cave system is about 560 km (348 miles) long. The **Sarawak Chamber in Malaysia is the world's biggest Cavern**. It is700 m (2,297 ft) long and over 70 m (230 ft) high.

Last Updated: 6-09-2002 Important Notices

Done Internet

Organised



reset

Thank you for your time and the valuable information you are submitting.

All of these questionnaires are translated to Arabic as mentioned in (Section 5.2.3) and they are illustrated in the following sections.

D.4 Colour Usability Questionnaire (Arabic version)

إستبانة لمعرفة الدور الذي تلعبه الألوان في الإنترنت تصميم صفحات

الصفحات تلعب دوراهاما في عملية عند تصفحك صفحات الإنترنت، هل لاحظت أن الألوان المستخدمة في هذه الإستبانة عبارة عن محاولة لإكتشاف العلاقة بين الألوان شد إنتباهك لهذه الصفحات أو تنفيرك منها. هذه الإنترنت و شد انتباه المتصفحين لها المستخدمة في صفحات

السؤال رقم 1

ما هـي لغتك الأم ؟

- عربي ○
- إنجليزي ○
- أخرى ○

السؤال رقم 2

أي مجموعات الألوان التالية تفضل؟

- النص الأزرق الغامق على الخلفية البيضاء
- الغامقة النص الأبيض على الخلفية الزرقاء
- النص الأسود على الخلفية الصفراء
- السوداء النص الأصفر على الخلفية

السؤال رقم 3

بلون مختلف عن بقية النص؟ هل تفضل أن تكون العناوين الفرعية و العبارات الهامة مكتوبة

- نعم
- لا

السؤال رقم 4

الصفحة أي أجزاء التصميم تفضل أن تكون بألوان مختلفة عن بقية محتويات

- العنوان الرئيسي للوثيقة
- العبارات الهامة
- الجداول
- العناوين الفرعية
- لا شئ

السؤال رقم 5

هل تفضل الصفحات ذات الخلفية الداكنة؟

- نعم
- لا

السؤال رقم 6

حدد الرمز الأقرب الذي يمثله كل من الألوان التاليه؟

- أبيض
- أسود
- احمر
- أزرق
- أصفر
- اخضر

السؤال رقم 7

بم تذكرك تركيبات الألوان التالية

أبيض لون الخلفية و **أحمر** لون النص

أبيض لون الخلفية و **أزرق** لون النص

أبيض لون الخلفية و **أخضر** لون النص

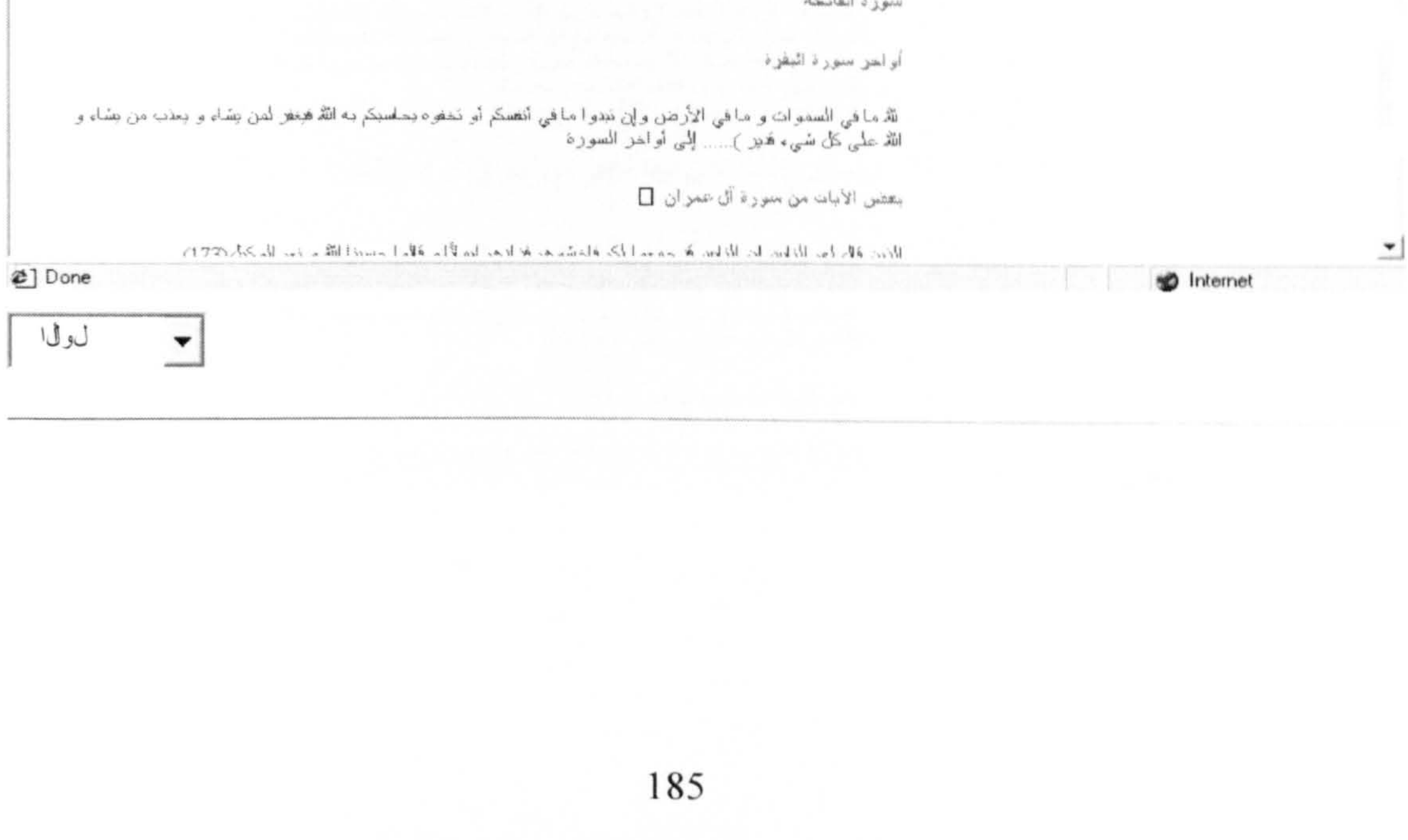
أسود لون الخلفية و **أصفر** لون النص

أحمر لون الخلفية و **أخضر** لون النص

من الأسئلة التاليه ملاحظة: سوف يتم إستخدام نفس الأربعة الأمثلة في كل

السؤال رقم 8

الأفضل رتب الأربعة الأمثلة التاليه حسب تفضيلك بحيث يكون الأول هو

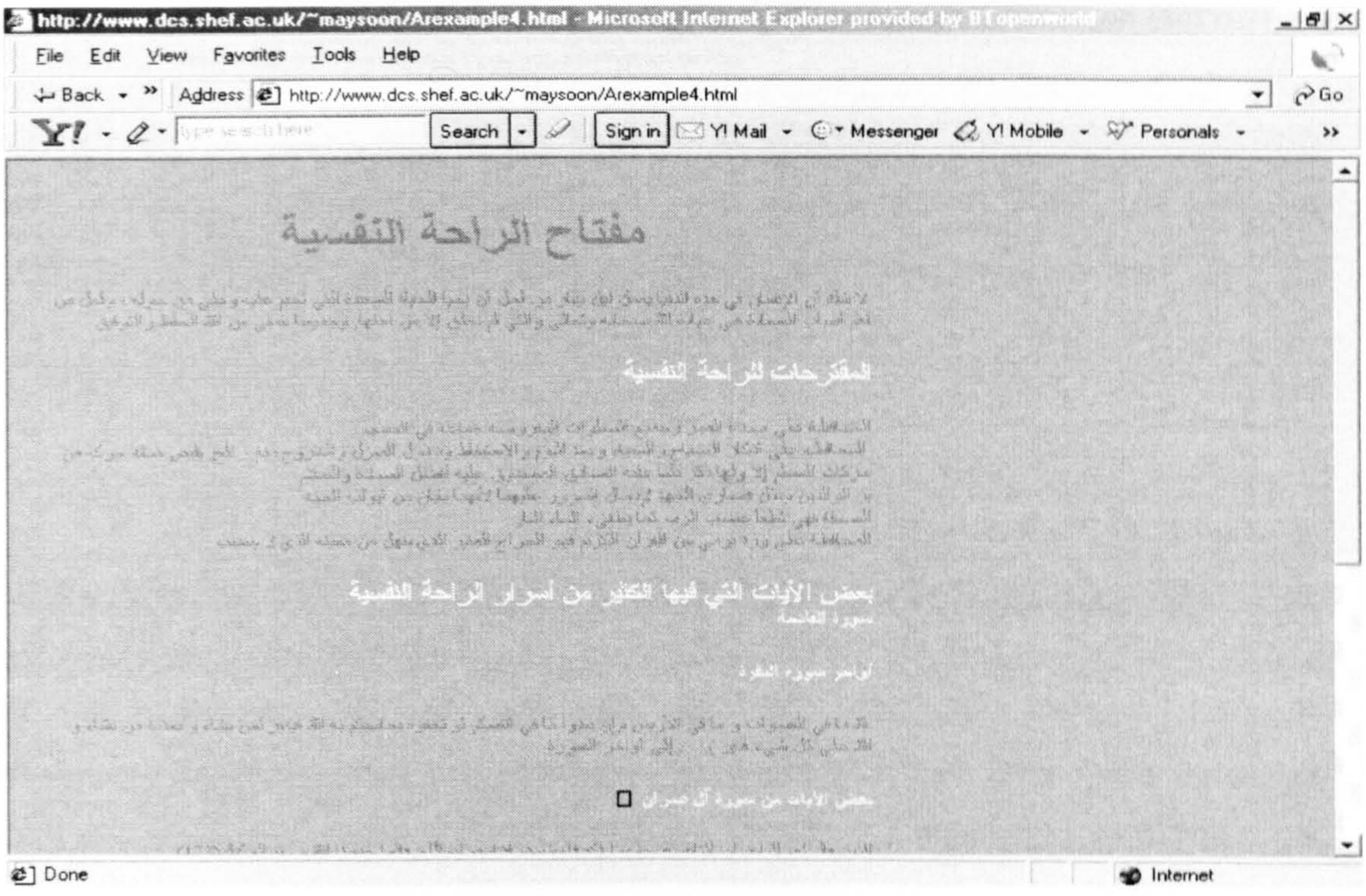


لولا


http://www.dcs.shef.ac.uk/~maysoon/Arexample2.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/Arexample2.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
بعض الآيات من سورة آل عمران
Done
Internet

لولا

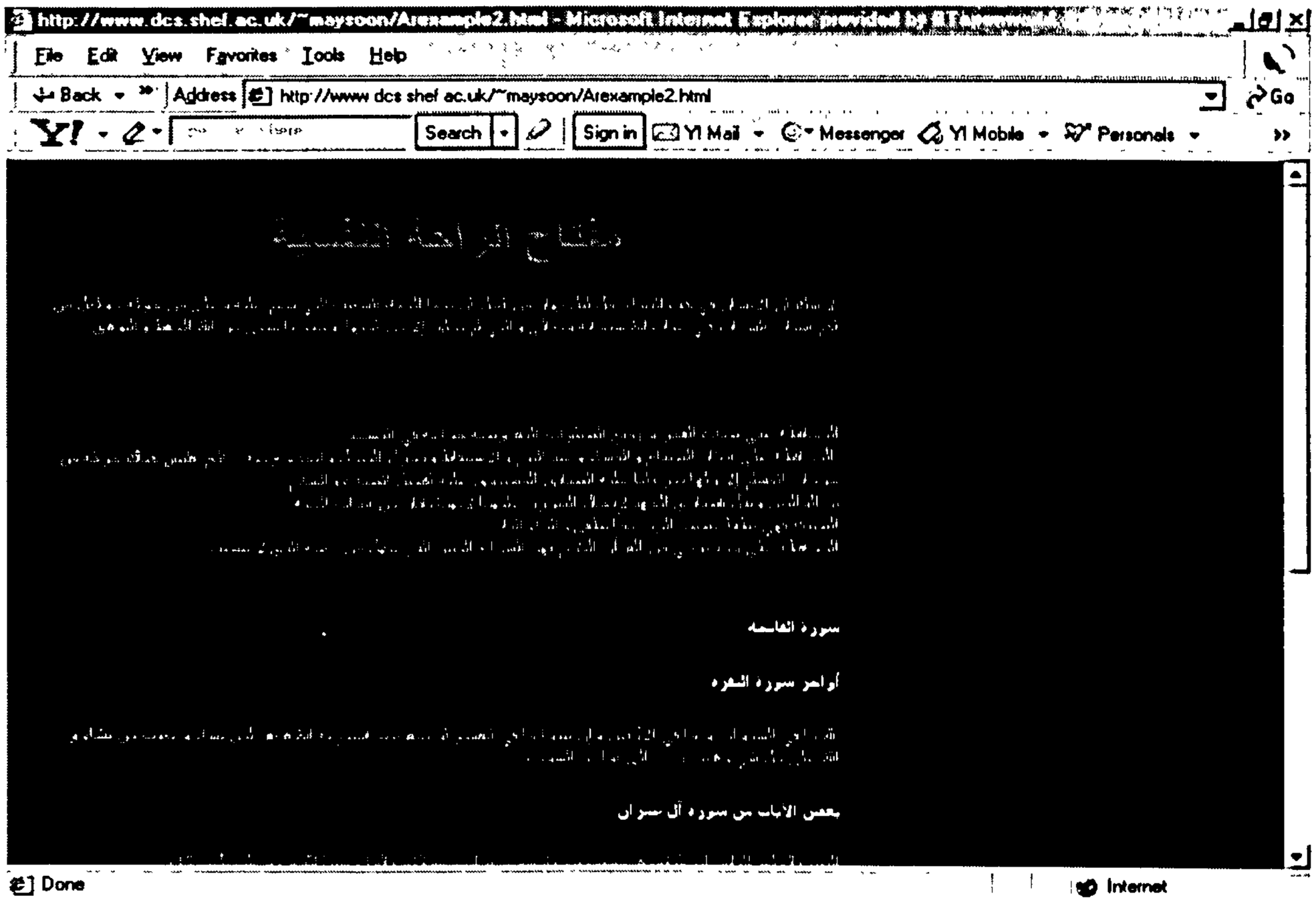

http://www.dcs.shef.ac.uk/~maysoon/Arexample3.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/Arexample3.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل لكل نهار من أجل أن يحيا الحياة السعيدة التي ينعم عليه وعلى من حوله . ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى و التي لم نخلق إلا من أجلها وجعلنا يتمنى من الله الحفظ و التوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح و المساء وعند النوم و الاستيقاظ ودخول المنزل والخروج منه... الخ فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ننهل من معينه الذي لا ينضب
بعض الأيات التي فيها الكثير من أسرار الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
لله ما في السموات و ما في الأرض وإن تبدوا ما في أنفسكم أو تخفوه يحاسبكم به الله فيغفر لمن يشاء و يعذب من يشاء و الله على كل شيء قدير) ..... إلى أواخر السورة
بعض الآيات من سورة آل عمران
Done
Internet



السؤال رقم 9

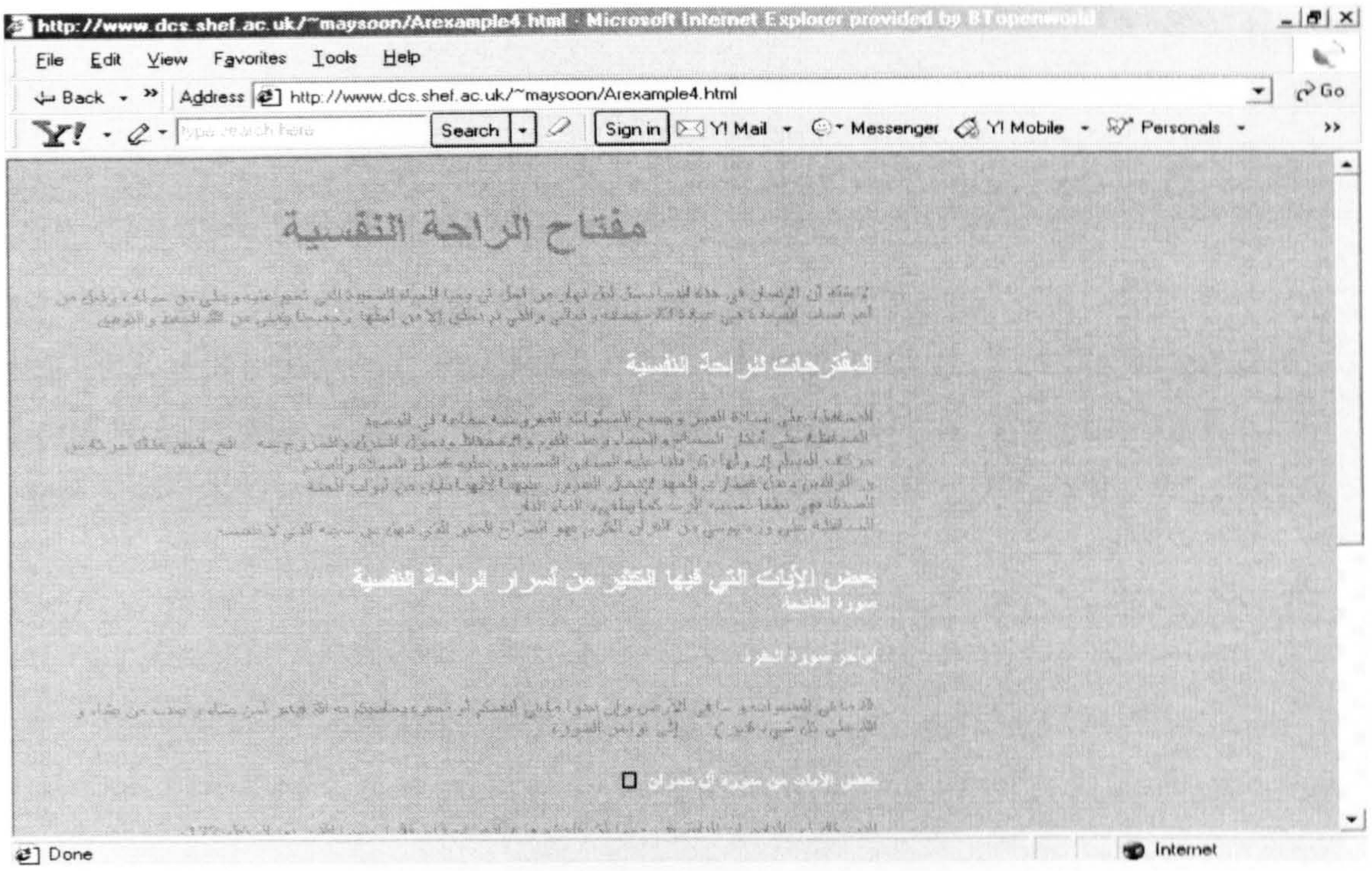
التالية حدد درجة جاذبية الألوان في كل من التصاميم الأربعة

جيد


http://www.dcs.shef.ac.uk/~maysoon/Arexample2.html - Microsoft Internet Explorer provided by IT
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/Arexample2.html
Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
بعض الآيات من سورة آل عمران
Done
Internet

ج ▼


http://www.dcs.shef.ac.uk/~maysoon/Arexample3.html - Microsoft Internet Explorer provided by IT
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/Arexample3.html
Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل لنيل شيء من أجل أن يحيا الحياة السعيدة التي ينعم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه . الخ فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ندخل من محبته الذي لا ينضب
بعض الأيات التي فيها الكثير من أسرار الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
لله ما في السموات و ما في الأرض وإن تبدوا ما في أنفسكم أو تخفوه يحاسبكم به الله فيغفر لمن يشاء و يعذب من يشاء و الله على كل شيء قدير) ..... إلى أواخر السورة
بعض الآيات من سورة آل عمران
Done
Internet



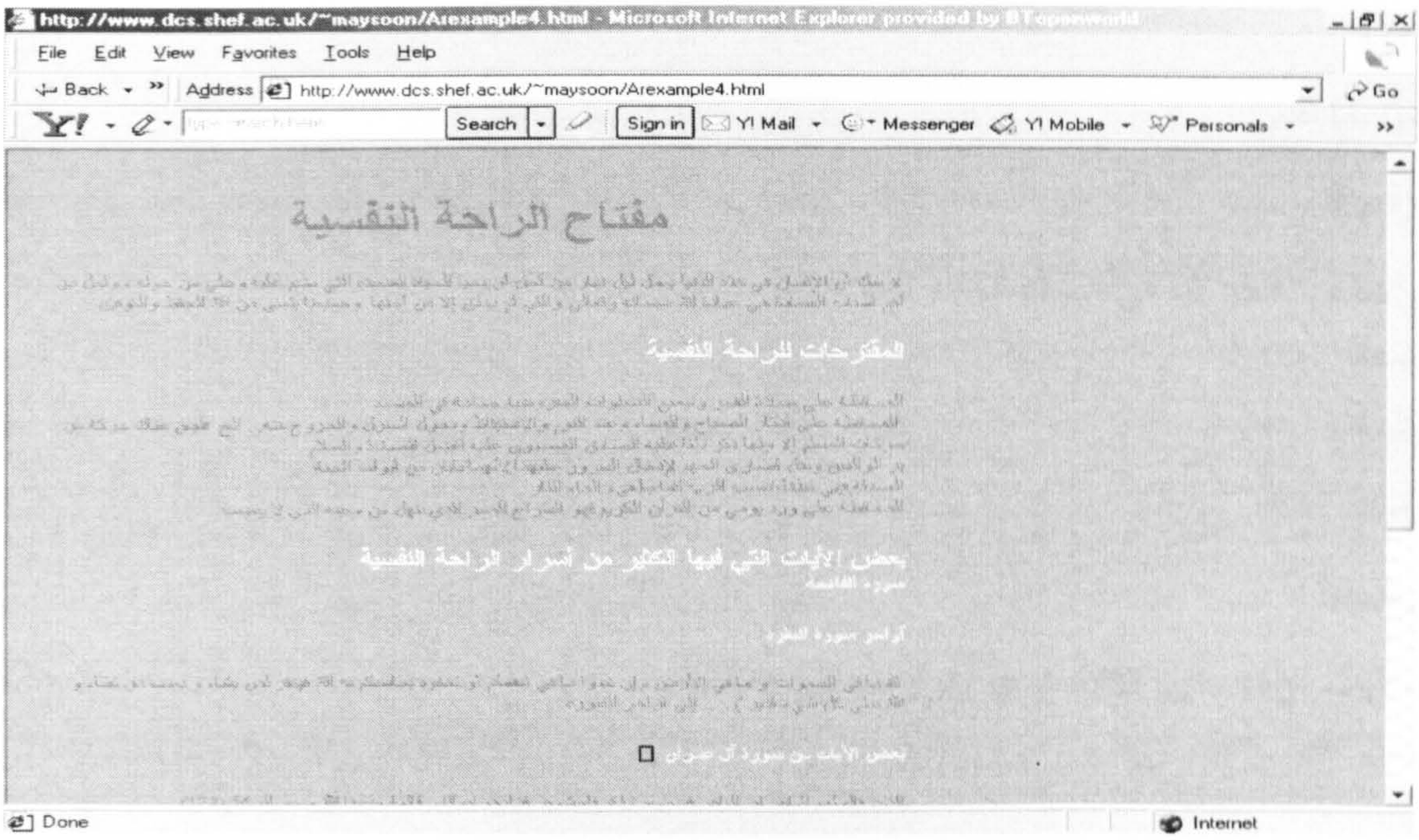
السؤال رقم 10

التالية حدد درجة جودة قراءة كل الألوان في التصاميم الأربعة

ي ج


http://www.dcs.shef.ac.uk/~maysoon/Arexample2.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back
Address http://www.dcs.shef.ac.uk/~maysoon/Arexample2.html
Go
Search
Sign in
Y! Mail
Messenger
Y! Mobile
Personals
مفتاح الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
بعض الآيات من سورة آل عمران
Done
Internet


http://www.dcs.shef.ac.uk/~maysoon/Arexample3.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back
Address http://www.dcs.shef.ac.uk/~maysoon/Arexample3.html
Go
Search
Sign in
Y! Mail
Messenger
Y! Mobile
Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي تخيم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى و التي لم نخلق إلا من أجلها وجميعنا يتمنى من الله الحفظ و التوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح و المساء وعند النوم و الاستيقاظ ودخول المنزل و الخروج منه ... الخ فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة و السلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ينهل من معجزته الذي لا ينضب
بعض الآيات التي فيها الكثير من أسرار الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
لله ما في السموات و ما في الأرض وإن تبدوا ما في أنفسكم أو تخفوه يحاسبكم به الله فيغفر لمن يشاء و يعذب من يشاء و الله على كل شيء قدير) ..... إلى أواخر السورة
بعض الآيات من سورة آل عمران
Done
Internet



ءاغل!

القيمة التي قدمتموها وفي الختام تتقدم الباحثة لكم بالشكر الجزيل على المعلومات

D.5 Text Usability Questionnaire (Arabic version)

إستبانة لمعرفة الدور الذي يلعبه تباين النصوص في تصميم صفحات الإنترنت

هذه الصفحات تلعب دورا هاما عند تصفحك صفحات الإنترنت، هل لاحظت أن مظهر النصوص المستخدمة في هذه الإستبانة عبارة عن محاولة لإكتشاف العلاقة بين .في عملية شد إنتباهك لهذه الصفحات أو تنفيرك منها صفحات الإنترنت و شد انتباه المتصفحين لها خصائص النصوص المستخدمة في

السؤال رقم 1

ما هو الجزء الذي استدعى إنتباهك أولا في هذه الصفحة؟

الجزء الأيسر العلوي ○

وسط الصفحة ○

وسط الصفحة ○

ذكره لا شئ مما تقدم ○

السؤال رقم 2

الجزء الذي تعمدت النظر إليه في هذه الصفحة أولا هو

شعار الجامعة ○

عنوان الوثيقة ○

الإستبانة أسئلة ○

السؤال رقم 3

المعلومات الهامة أي الأساليب التالية تفضل لإبراز العناوين الفرعية و

خط عريض ○

خط مائل ○

خط مسطر ○

خط كبير ○

مختلف خط بلون ○

السؤال رقم 4

سواها أي العناوين الفرعية التالية يمكن قرائتها بشكل أسرع من

عنوان عريض ○

عنوان مائل ○

عنوان مسطر ○

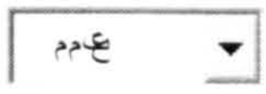
عنوان بخط كبير ○

مختلف عنوان بلون ○

السؤال رقم 5

جودة القراءة في الإعتبار حدد مستوى متعة قراءة كل من النصوص الثلاثة التالية مع وضع


http://www.dcs.shef.ac.uk/~maysoon/Aremphasizeexample1.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/Aremphasizeexample1.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي تخيم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها. وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه...
الخ. فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ننهل من معينه الذي لا ينضب
بعض الأيات التي فيها الكثير من أسرار الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
Done
Internet


عام


http://www.dcs.shef.ac.uk/~maysoon/Aremphasizeexample2.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/Aremphasizeexample2.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي تخيم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها .
وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه ... الخ.
فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ننهل من معينه الذي لا ينضب
بعض الأيات التي فيها الكثير من أسرار الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
لله ما في السموات و ما في الأرض وإن تبدوا ما في أنفسكم أو تخفوه يحاسبكم به الله فيغفر لمن يشاء و يعذب من يشاء و الله على كل شيء قدير).....إلى أواخر السورة
Done
Internet

السؤال رقم 6

قيم كل من النصوص الأربعة التالية بحسب سرعة تعرفك علىمضمونها

لم

ل هـ

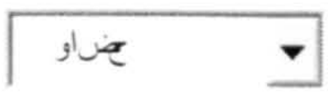

http://www.dcs.shef.ac.uk/~maysoon/ArfScanningexample2.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/ArfScanningexample2.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي تخيم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على الأذكار في الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه... الخ فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ننهل من معينه الذي لا ينضب
Last Updated: 27-01-2003
Important Notices
Done
Internet

ل هـ


http://www.dcs.shef.ac.uk/~maysoon/ArfScanningexample3.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/ArfScanningexample3.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي تخيم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على الأذكار في الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه... الخ فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ننهل من معينه الذي لا ينضب
Last Updated: 27-01-2003
Important Notices
Done
Internet

السؤال رقم 7

حدد مستوى جودة وضوح قراءة كل مما يلي



حضاو


http://www.dcs.shef.ac.uk/~maysoon/ArfontSizeexample2.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/ArfontSizeexample2.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي تخيم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه... الخ فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ننهل من معينه الذي لا ينضب
بعض الآيات التي فيها الكثير من أسرار الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
لله ما في السموات و ما في الأرض وإن تبدوا ما في أنفسكم أو تخفوه يحاسبكم به الله فيغفر لمن يشاء و يعذب من يشاء و الله على كل شيء قدير)..... إلى أواخر السورة
بعض الآيات من سورة آل عمران
Done
Internet

حضاو


http://www.dcs.shef.ac.uk/~maysoon/ArfontSizeexample3.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/ArfontSizeexample3.html Go
Search Sign in Y! Mail Messenger Y! Mobile Personals
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي تخيم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها. وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه...
Done
Internet

عاغل!

القيمة التي قدمتموها وفي الختام تتقدم الباحثة لكم بالشكر الجزيل على المعلومات

D.6 Title Usability Questionnaire (Arabic version)

إستبانة لمعرفة الدور الذي يلعبه ترتيب العناوين في الإنترنت تصميم صفحات

الأسطر المستخدمة في هذه الصفحات عند تصفحك صفحات الإنترنت، هل لاحظت أن عنوان الوثيقة و بداية الصفحات أو تنفيرك منها. هذه الإستبانة عبارة عن محاولة لإكتشاف تلعب دوراً هاماً في عملية شد إنتباهك لهذه العناوين المستخدمة في صفحات الإنترنت و شد انتباه المتصفحين لها العلاقة بين خصائص

السؤال رقم 1

حدد عنوان هذه الصفحة حسب إعتقادك

- الإنترنت إستبانة لمعرفة الدور الذي يلعبه ترتيب العناوين في تصميم صفحات
- Research Usability Questionnaire

السؤال رقم 2

هل تفضل وجود عدد كبير من الروابط في صفحة الإنترنت؟

السؤال رقم 2

هل تفضل وجود عدد كبير من الروابط في صفحة الإنترنت؟

نعم ○

لا ○

السؤال رقم 3

قيم العناوين التالية من حيث سهولة تذكر كل منها؟

ل هـ

السؤال رقم 4

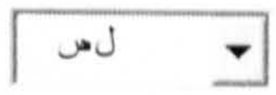
قيم جودة تصميم الصفحات الثلاث التالية ؟

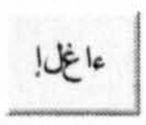

http://www.dcs.shef.ac.uk/~maysoon/ArAlignmentExample1.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/ArAlignmentExample1.html
Go
Search Y! Mail Personals Sign in
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي ينعم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها. وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه... الخ فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ننهل من معينه الذي لا ينضب
بعض الأيات التي فيها الكثير من أسرار الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
لله ما في السموات و ما في الأرض وإن تبدوا ما في أنفسكم أو تخفوه يحاسبكم به الله فيغفر لمن يشاء و يعذب من يشاء و الله على كل شيء قدير)..... إلى أواخر السورة
بعض الآيات من سورة آل عمران □
Done
Internet


http://www.dcs.shef.ac.uk/~maysoon/ArAlignmentExample3.html - Microsoft Internet Explorer provided by BTopenworld
File Edit View Favorites Tools Help
Back Address http://www.dcs.shef.ac.uk/~maysoon/ArAlignmentExample3.html
Go
Search Y! Mail Personals Sign in
مفتاح الراحة النفسية
لا شك أن الإنسان في هذه الدنيا يعمل ليل نهار من أجل أن يحيا الحياة السعيدة التي ينعم عليه وعلى من حوله ، ولعل من أهم أسباب السعادة هي عبادة الله سبحانه وتعالى والتي لم نخلق إلا من أجلها. وجميعنا يتمنى من الله الحفظ والتوفيق
المقترحات للراحة النفسية
المحافظة على صلاة الفجر وجميع الصلوات المفروضة جماعة في المسجد
المحافظة على أذكار الصباح والمساء وعند النوم والاستيقاظ ودخول المنزل والخروج منه... الخ فليس هناك حركة من حركات المسلم إلا ولها ذكر دلنا عليه الصادق المصدوق عليه أفضل الصلاة والسلام
بر الوالدين وبذل قصارى الجهد لإدخال السرور عليهما لأنهما بابان من أبواب الجنة
الصدقة فهي تطفأ غضب الرب كما يطفيء الماء النار
المحافظة على ورد يومي من القرآن الكريم فهو السراج المنير الذي ننهل من معينه الذي لا ينضب
بعض الأيات التي فيها الكثير من أسرار الراحة النفسية
سورة الفاتحة
أواخر سورة البقرة
لله ما في السموات و ما في الأرض وإن تبدوا ما في أنفسكم أو تخفوه يحاسبكم به الله فيغفر لمن يشاء و يعذب من يشاء و الله على كل شيء قدير)..... إلى أواخر السورة
□
Done
Internet

السؤال رقم 5

في الإعتبار؟ حدد مستوى سهولة قراءة الصفحات التالية، واضعاً بداية السطر





القيمة التي قدمتموها وفي الختام تتقدم الباحثة لكم بالشكر الجزيل على المعلومات

Bibliography

Abulkhair, M. and S. North (2003). A Fuzzy Model to Measure Colour Contrast as an Aspect of Web Usability. Human - Computer Interaction, Crete, Greece, Lawrence Erlbaum Associates. Eds. Jacko J. and Stephanidis C. 1(4): 733-737.

Aljazeera (2003). Aljazeera Web Page, http://www.aljazeera.net/.

Allen, R. (1984). Pocket Oxford Dictionary. Oxford, Clarendon Press.

Amant, R. (2000). "Interface Agents as Surrogate Users." Intelligence: 29 - 38.

Amant, R. and L. Zettlemoyer (2000). The User Interface as an Agent Environment. International Conference on Autonomous Agents, Barcelona, Spain, ACM Press : 483- 490.

Anthony, M. and P. Bartlett (1999). Neural Network Learning: Theoretical Foundations. Cambridge, UK, Cambridge University Press.

Antonacopoulos, A. and D. Karatzas (2002). Fuzzy Segmentation of Characters in Web Images Based on Human Colour Perception. Document Analysis Systems, Princeton, NY, Springer-Verlag Heidelberg. Eds. D. Lopresti, J. Hu and R. Kashi : 295-306.

Arabic Color Usability Questionnaire (2003). http://www.dcs.shef.ac.uk/~maysoon/ArabicColourQuestionnaire.html

Arabic Text Usability Questionnaire (2003). http://www.dcs.shef.ac.uk/~maysoon/ArabicTextQuestionnaire.html

Arabic Title Usability Questionnaire (2003). http://www.dcs.shef.ac.uk/~maysoon/ArabicTitleQuestionnaire.html

Arnold, K., J. Gosling, et al. (2000). The Java Programing Language. Boston, Addison- Wesley.

Badre, A. (2002). Shaping Web Usability: Interaction Design in Context. Boston, USA, Addision-Wesley.

Bank, A. (2003). Alahli Bank Web Page. http://www.alahli.com/arabic/.

Barber, K. S. and C. E. Martin (2001). Dynamic reorganization of decision-making groups. International Conference on Autonomous Agents, Montreal, Quebec, Canada, ACM Press : 513 – 520.

Bayardo, R., W. Bohrer, et al., Eds. (1998). InfoSleuth: agent-based Semantic Integration of Information in Open and Dynamic environment. Readings in Agents. San Francisco, Morgan Kaufmann Publisher, INC.

BBC (2003). BBC Web Page, http://www.bbc.co.uk/.

Beale, R. and A. Wood (1994). Agent-based Interaction. Human Computer Interaction, Glasgow, Cambridge University Press New York: 239- 245.

Becker, S. A. (2002). "An Explortory Study on Web Usability and the Internationalization of US E-Businesses." Journal of Electronic Commerce Research 3(4): 265-278.

Bevan, N. (1998). Usability Issues in Web Site Design. Usability Professionals Association Conference, Washigton DC, http://www.usability.serco.com/papers/usweb98.pdf (3): 22- 26.

Bevan, N. (1999). Design for Usability. Human Computer Interaction International, Munich. Lawrence Erlbaum Associate, Eds. Bullinger H. J., Ziegler, J.: 22- 27.

Bevan, N., J. Kirakowski, et al. (1991). What is Usability? Proceedings of the 4th International Conference on Human Computer Interaction, Stuttgart. Ed. Bullinger H. J.: 651-655.

Bevan, N. and M. Macleod (1994). "Usability Measurement in Context." Behaviour and Information Technology 13(1&2): 132- 145.

Blumberg, B., M. Downie, et al. (2002). Integrated Learning For Interactive Synthetic Characters. International Conference on Computer Graphics and Interactive Techniques, San Antonio, Texas, ACM Press : 417- 426.

Borges, J., I. Morales, et al. (1996). Guidelines for Designing Usable World Wide Web Pages. Human Factors in Computing Systems, Vancouver, British Columbia, Canada, ACM Press : 277- 278.

Brinck, T. and E. Hofer (2002). Automatically evaluating the Usability of Web Sites. Human Factors in Computing Systems, Minneapolis, Minnesota, ACM, Press: 906- 907.

Brown, S., E. S. Jr., et al. (1998). Unsing Explicit Requirements and Metrics for Interface Agent User Model Correction. Autonomous Agent, Minneapolis, MN, ACM Press : 1- 7.

Bryman, A. (2001). Social Research Methods. Oxford, Oxford University Press.

Burmeister, O. K. (2000). Usability testing: revisiting informed consent procedures for testing internet sites. ACM International Conference Proceeding Series Selected papers from the second Australian Institute conference on Computer ethics, Canberra, Australia, Australian Computer Society, Inc 1: 1- 7.

Burns, R. (2000). Introduction to Research Methods. London, SAGE Publications Ltd.

Calongne, C. M. (2001). "Designing for Web Site Usability." The Journal of Computing in Small Colleges 16(3): 39 - 45.

Canfora, G. and L. Troiano (2002). The importance of dealing with uncertainty in the evaluation of software engineering methods and tools. international conference on Software engineering and knowledge engineering, Ischia, Italy, ACM Press : 691 - 698.

Caudill, M. and C. Butler (1990). Naturally Intelligent Systems. Cambridge, Massachusetts, The MIT Press.

Cesta, A. and D. D'Aloisi (1996). "Building interfaces as personal agents: a case study." ACM SIGCHI Bulletin 28(3): 108 - 113.

Chadwick-Dias, A., M. McNulty, et al. (2003). Web usability and age: how design changes can improve performance. ACM Conference on Universal Usability, Vancouver, British Columbia, Canada, ACM Press : 30 - 37.

Chi, E. H., P. Pirolli, et al. (2001). Using information scent to model user information needs and actions and the Web. Human Factors and Computing Systems, Seattle, Washington, ACM Press : 490 - 497.

Chi, E. H., A. Rosien, et al. (2003). The Bloodhound Project: Automating Discovery of Web Usability Issues Using the InfoScent Simulator. Human Factors in Computing Systems, Ft. Lauderdale, Flourida, ACM Press 1(5) : 505-512.

Coen, M. (1994). SodaBot: A Software Agent Environment and Construction System. Cambridge, MA, MIT AI Lab Technical Report 1493.

Cohen, P., A. Cheyer, et al., Eds. (1998). An Open Agent Architecture. Readings in Agents. San Francisco, Morgan Kaufmann Publisher, INC.

Crescent, R. (2003). Red Crescent Web Page. http://www.srcs.org.sa/.

Cross, R. (2003). Red Cross Web Page. http://www.redcross.org.uk/homepage.asp.

Czinkota, M. and I. Ronkainen (1996). Global Marketing. New York, The Dryden Press.

Das, S., A. Caglayan, et al. (1998). Increasing agent autonomy in dynamic environments. International Conference on Autonomous Agents, Minneapolis, Minnesota, ACM Press : 309 - 316.

Dictionary, A. H. (1994). The American Heritage Dictionary. New York, Dell Publishing.

Dix, A., J. Finlay, et al. (1998). Human-Computer Interaction. Upper Saddle River, NJ, Prentice Hall.

Do, O., E. March, et al. (1996). Intelligent Agents & the Internet: Effects On Electronic Commerce and Marketing, University of Nevada, Reno. http://bold.coba.unr.edu/odie/paper.html

El-Nasr, M. S., J. Yen, et al. (2000). "FLAME - Fuzzy Logic Adaptive Model of Emotions." Automous Agents and Multi-agent Systems 3: 219-257.

English Color Usability Questionnaire (2003). http://www.dcs.shef.ac.uk/~maysoon/EnglishColourQuestionnaire.html

English Text Usability Questionnaire (2003). http://www.dcs.shef.ac.uk/~maysoon/EnglishTextQuestionnaire.html

English Title Usability Questionnaire (2003). http://www.dcs.shef.ac.uk/~maysoon/EnglishTitleQuestionnaire.html

Etgen, M. and J. Cantor (1999). What does getting WET(Web Event-logging Tool) mean for web usability? Human Factor and the Web, Gaithersburg, MD.

Etizion, O. and D. Weld, Eds. (1998). A Softbot-Based Interface to the Internet. Readings in Agents. San Francisco, Morgan Kaufmann Publisher, INC.

Foner, L., Ed. (1993). What's an Agent, Anyway? A Sociological Case Study. Agents Memo 93-01. Cambridge, MA, MIT Media Lab.

Forrest, S. (1996). "Genetic algorithms." ACM Computing Surveys (CSUR) 28(1): 77 - 80.

Franklin, S. and A. Graesser (1996). Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. Third International Workshop on Agent Theories, Architectures, and Languages, Berlin Heidelberg, Germany, Springer-Verlag : 21-35.

Friedman, B. and H. Nissenbaum (1996). "Bias in computer systems." ACM Transactions on Information Systems 14(3): 330-347.

Fujishima, Y. (1997). An Interface Agent for Nonroutine Tasks. Interactive User Interfaces, Orlando, Florida, ACM Press : 213- 216.

Gee, K. (2001). "The Ergonomics of Hypertext Narrative: Usability Testing as a Tool or Evaluation and Redesign." ACM Journal of Computer Documentation 25(1): 3- 16.

Genesereth, M. (1994). "Software Agents." Communications of the ACM 37(7): 48-53.

Gmytrasiewicz, P. and M. Huhns (2000). "The Emergence of Language Among Autonomous Agents." IEEE Internet Computing: 90-92.

Goecks, J. and J. Shavlik (2000). Learning Users interests by unobtrusively observing their normal behaviour. Interactive User Interfaces, New Orleans, ACM Press: 129- 132.

Goonetilleke, R., C. Duffy, et al. (2001). Customer needs in Web Based Interaction: A Macro view of Usability. Human - Computer Interaction : 365-369.

Goren-Bar, D. and T. Kuflik (2004). Don't miss-r --: recommending restaurants through an adaptive mobile system. International Conference on Intelligent User Interfaces, Funchal, Madeira, Portugal, ACM Press : 250 - 252.

Graham, I. (2003). A Pattern Language for Web Usability. London, Addison - Wesley.

Hackos, J. and J. Redish (1998). User and Task Analysis for Interface Design. New York, John Wiley & Sons, Inc.

Harris, C., C. Moore, et al. (1993). Intelligent Control Aspects of Fuzzy Logic and Neural Nets. Singapore, World Scientific.

Hawthorn, D. (2003). How universal is good design for older users? ACM Conference on Universal Usability, Vancouver, British Columbia, Canada, ACM Press : 38- 45.

Head, A. (1999). Design Wise: a guide for evaluating the interface design of information resources. Medford, NJ, Information Today.

Hertz, J., A. Krogh, et al. (1991). Introduction to the Theory of Neural Computation. Redwood City, California, Addison-Wesely Publishing Company.

Hilbert, D. and D. Redmiles (2000). "Extracting usability information from user interface events." ACM Computing Surveys (CSUR) 32(4): 384 - 421.

Hong, J. and J. Landay (2001). WebQuilt: a Framework for Capturing and Visualizing the Web Experience. international conference on World Wide Web, Hong Kong, Hong Kong: 717-724.

Hong, J. I., J. Heer, et al. (2001). "WebQuilt: A proxy-based approach to remote web usability testing." ACM Transactions on Information Systems 19(3): 263-285.

Huhns, M. (2000). "An Agent-Based Global Economy." IEEE Internet Computing: 83-84.

Huhns, M. and M. Singh (1997). "Ontologies for Agents." IEEE Internet Computing 1(6): 81-83.

Huhns, M. and M. Singh, Eds. (1998). Readings in Agents. San Francisco, Morgan Kaufmann Publisher, INC.

IDC (2001). Global Internet Population, CyberAtlas, http://cyberatlas.internet.com/big_picture/geographics/print/0,,5911_558061,00.html

Isbell, C., C. Shelton, et al. (2001). A Social Reinforcement Learning Agent. Autonomous Agents and Multiagents, Montereal, Quebec, Canada, ACM Press : 377 - 384.

ISO (1991). Ergonomic Requirements for Office Work with Visual Display Terminals, International Standard Organisation ISO 9241.

ISO (1991). Software Product Evaluation - Quality Charactristics and Guidelines for their Use, International Standard Organisation ISO DIS 9126.

Ivory, M. (2000). Web TANGO: towards automated comparison of information-centric web site designs. Human Factors in Computing Systems, The Hague, The Netherlands, ACM Press : 329 - 330.

Ivory, M. (2001). An Empirical Foundation for Automated Web Interface Evaluation. Computer Science. Berkeley, University of California at Berkeley.

Ivory, M. (2003). Characteristics of Web Site Designs: Reality vs. Recommendation. Human - Computer Interaction, Crete, Greece, Lawrence Erlbaum Associates. Eds. Jacko J. and Stephanidis C. 1(4): 773-777.

Ivory, M. and M. Hearst (2001). "The State of the Art in Automating Usability Evaluation of User Interfaces." ACM Computing Surveys (CSUR) 33(4): 470-516.

Ivory, M. and M. Hearst (2002). Automated Usability Evaluation of Web Interfaces. Human factors in Computing Systems Workshop, Minneapolis, Minnesota, ACM, Press, http://www.usabilityfirst.com/auto-evaluation/paper_Ivory_Hearst.html

Ivory, M. and M. Hearst (2002). "Improving Web Site Design." IEEE Internet Computing: 56-63.

Ivory, M. and M. Hearst (2002). Statistical Profiles of Highly-Rated Web Sites. Human Factors in Computing Systems, Minneapolis, Minnesota, USA, ACM Press: 367 – 374.

Ivory, M. and M. Hearst (2002). "Towards Quality Checkers for Web Site Designs." IEEE Internet Computing 6(2): 56- 63.

Ivory, M., R. Sinha, et al. (2001). Empirically validated web page design metrics. Human Factors in Computing Systems, Seattle, Washington, ACM Press: 53 – 60.

John, B. E. (1996). "Evaluating Usability Evaluation Techniques." ACM Computing Surveys (CSUR) 28(4es).

Jones, G. M. (2000). Usability services at the University of Maryland: who, what and how. User Services Conference, Richmond, Virginia, ACM Press: 343 – 349.

Junhong Nie, D. L. (1995). Fuzzy-Neural Control Principles, Algorithms and Applications. New York, Prentice Hall.

Karis, D. (2002). A proposal for the CHI'2002 workshop: Automatically Evaluating the usability of web sites. Human factors in Computing Systems Workshop, Minneapolis, Minnesota USA, ACM Press, http://www.usabilityfirst.com/auto-evaluation/

Kautz, H., B. Selman, et al. (1994). "Bottom-Up Design of Software Agents." Communications of the ACM 37(7): 143-146.

Keevil, B. (1998). Measuring the Usability Index of Your Web Site. The 16th annual International Conference on Computer Documentation, Quebec, Quebec, Canada, ACM Press: 271- 277.

Kerr, M. (2001). Tips and Tricks for Website Mangers. London, United Kingdom, Aslib-IMI.

Knapik, M. and J. Johnson (1998). Developing Intelligent Agents for Distributed Systems. New York, McGraw-Hill.

Koda, T. and P. Maes (1996). Agents with Faces: The Effects of Personification of Agents. Human Computer Interaction, London, The British HCI Group: 98-103.

Koyani, S. and S. Allison (2003). Use of research-based guidelines in the development of websites. Human Factors in Computing Systems, Ft. Lauderdale, Florida, ACM Press: 696 – 697.

Laarni, J. (2003). Effects of Color, Font Type and Word Style on User Preferences and Emotions. Human Computer Interaction International, Crete, Greece, Crete University Press, Ed. Stephanidis, C.: 31- 32.

Lange, D. and M. Oshima (1998). Programming and Deploying Java Mobile Agents with Aglets. Reading, MA, Addison- Wesley.

Lashkari, Y., M. Metral, et al., Eds. (1998). Collabrative Interface Agents. Readings in Agents. San Francisco, Morgan Kufmann Publisher, INC.

Laskowski, S., J. Landay, et al. (2002). Automatic Capture, Representation, and Analysis of User Behavior. Human Factors in Computing Systems, Minneapolis, Minnesota: 922- 923.

Laurel, B., Ed. (1990). The Art of Human-Computer Interface Design. New York, Addison-Wesley.

Law, L.-C. and E. Hvannberg (2002). Complementarity and Convergence of Heuristic Evaluation and Usability Test: A Case Study of UNIVERSAL Brokerage Platform. NoridiCHI, Arhus, Denmark, ACM press: 71-80.

Lee, K.-P. (1999). A Study on the Cultural Effects on User Interface Design with the Emphasis on the Cross-Cultural Usability Testing Through World Wide Web. Asian Design Conference ADC, Nagaoka-shi, Japan: 369 – 381.

Lengel, J. (2002). The Web Wizard's Guide to Web Design. Boston, Addison - Wesley's.

Lieberman, H. (1995). "Letizia: An Agent That Assists Web Browsing." IJCAI 1: 924-929.

Lieberman, H. (1997). Autonomous Interface Agents. Human Factors in Computing Systems, Atlanta, Georgia, ACM Press: 67-74.

Lieberman, H. (1997). Integrating user interface agents with conventional applications. Intelligent User Interfaces, San Francisco, California, ACM Press: 39 – 46.

Lieberman, H., B. Nardi, et al. (1999). Training agents to recognize text by example. International Conference on Autonomous Agents, Seattle, Washington, ACM Press: 116 – 122.

Lieberman, H., F. Paternò, et al. (2003). Perspectives on end user development. Human Factors and Computing Systems, Ft. Lauderdale, Florida, ACM Press: 1048 – 1049.

LIFT (2000). LIFT Onsite, UsableNet, Inc. http://www.usablenet.com/products_services/products_services.html.

Lin, C. (1994). Neural Fuzzy Control Systems with Structure and Parameter Learning. Singapore, World Scientific.

Lin, J., M. Newman, et al. (2000). DENIM: Finding A Tighter Fit Between Tools and Practice for web Site Design. Human Factors in Computing Systems, Hague, The Netherlands, ACM, Press: 510 – 517.

Lisboa, P., Ed. (1992). Neural Networks Current Applications. London, Chapman and Hall.

Little, J., M. Granger, et al. (2000). Integrating Cultural Issues into the Computer and Information Technology Curriculum. Helsinki, Finland: 136-154.

Lynch, P. and S. Horton (2001). Web Style Guide Basic Design Principles for Creating Web. New Haven, Yale University Press.

Lynden, S. and O. Rana (2002). Coordinating Learning Agents via Utility Assignment. Intelligent Data Engineering and Automated Learning, Manchester, UK, Springer- Verlag. Eds. Yin, H. Allinson, N. Freeman R. Keane J. Hubbard, S.: 219-224.

Maes, P. (1994). "Agents that reduce work and information overload." Communications of the ACM 37(7): 30 - 40.

Maes, P. (1995). "Artificial life meets entertainment: lifelike autonomous agents." Communications of the ACM 38(11): 108 - 114.

Maes, P. (1997). Intelligent Software. Intelligent User Interfaces, Orlando, FL, ACM Press: 41- 43.

Maes, P. and B. Shneiderman (1997). "Direct manipulation vs. interface agents." Interactions 4(6): 42 - 61.

Maes, P., B. Shneiderman, et al. (1997). Intelligent Software Agents vs. User-Controlled Direct Manipulation: A Debate. Human Factors and Computing Systems, Atlanta, Georgia, ACM Press: 105- 106.

Maes, P. and A. Wexelblat (1996). Interface Agents. Human Factors in Computing Systems, Vancouver, British Columbia, Canada, ACM Press: 369 – 370.

Maguire, M. (2000). Increasing the Influence of Usability Practices within the Design Process. Human Factors in Computing Systems, The Hague, The Netherlands, ACM Press: 305 – 305.

Maguire, M. and A. Dillon (1993). Usability measurement — its practical value to the computer industry. Human Factors in Computing Systems, Amsterdam, The Netherlands, Addison-Wesley Longman Publishing Co., Inc.

Mail, R. (2003). Royal Mail Web Page. http://www.royalmail.com/.

Manley, B. (2001). quoted in Tips and Tricks for Website Mangers. M. Kerr. London, Aslib-IMI.

Marcus, A. (2001). Cross-cultural user-interface design for work, home, play, and on the way. The 19th annual international conference on Computer documentation, Sante Fe, New Mexico, ACM Press: 221 – 222.

Marcus, A. and E. W. Gould (2000). "Crosscurrents Cultural Dimentions and Global Web User-Interface Design." Interactions, ACM: 32-46.

Meier, B. (1988). ACE: A Color Expert System for User Interface Design. Symposium on User Interface Software and Technology, Alberta, Canada, ACM Press: 117-128.

Melanie, M. (1996). An introduction to genetic algorithms. Cambridge, Mass, MIT Press.

Millier, M. (1997). Software Agents. Human Factors and Computing Systems, Atlanta, Georgia, ACM, Press: 194-195.

Mooney, M. (1998). Managing color in interactive systems. Human Factors and Computing Systems, Los Angeles, California, ACM Press: 169 – 170.

Moran, D., A. Cheyer, et al. (1997). Multimodal User Interfaces in the Open Agent Architecture. Intelligent User Interfaces, Orlando, Florida, ACM Press: 61-68.

Mukaidono, M. (2001). Fuzzy Logic for Beginners. Singapore, World Scientific Publishing Co. Pte. Ltd.

Murch, R. and T. Johnson (1999). Intelligent Software Agents. Upper Saddle River, New Jersey, Prentice Hall.

Nassif, A. (2003). A Cross-National Comparison of Saudi and British Televesion Advertisments. Sheffield, The University of Sheffield.

Nations, U. (1999). United Nations Nations Unies. http://www.un.org/Depts/OHRM/brochure.htm.

Neville, M. and A. Sibley (2002). Developing a generic genetic algorithm. Annual International Conference on Ada, Houston, Texas, ACM Press: 45 – 52.

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. Conference on Human Factors and Computing Systems, Monterey, California, ACM Press: 373 – 380.

Nielsen, J. (1993). Usability Engineering. San Francisco, California, Morgan Kaufmann Publishers, Inc.

Nielsen, J. (1994). Usability Inspection Methods. Human Factors in Computing Systems, Boston, ACM Press: 413 – 414.

Nielsen, J. (1995). Usability inspection methods. Human Factors in Computing Systems, Denver, ACM Press: 377 – 378.

Nielsen, J. (2000). Designing Web Usability: The Practice of Simplicity. Indianapolis, USA, New Rider publishing.

Nielsen, J. (2002). Top Ten Guidelines for Homepage Usability, Alertbox. http://www.useit.com/alertbox/20020512.html.

Nielsen, J. (2003). The Ten Most Violated Homepage Design Guidelines, Alertbox. http://www.useit.com/alertbox/2002052.html.

Nielsen, J. (2003). Top Ten Web Design Mistakes of 2003, Alertbox. http://www.useit.com.alertbox/20020512.html.

Nielsen, J. (2004). Ten Steps for Cleaning Up Information Pollution, Alertbox. http://www.useit.com/alertbox/20020512.html.

Nielsen, J. and T. Landauer (1993). A mathematical model of the finding of Usability problems. Human Factors in Computing Systems, Amsterdam, The Netherlands, Addison-Wesley Longman Publishing Co., Inc.

Nielsen, J. and M. Tahir (2002). Homepage Usability 50 Websites Deconstructed. Indianapolis, Ind., New Rider publishing.

Nielsen, J., C. Torkil, et al. (2002). Getting Access to What Goes on in People's Heads? Reflections on the Think-aloud Technique. NordiCHI, Arhus, Denmark, ACM, Press: 101- 111.

Nodder, C., G. Williams, et al. (1999). Evaluating the usability of an evolving collaborative product —changes in user type, tasks and evaluation methods over time. Conference on Supporting Group Work, Phoenix, Arizona, ACM Press: 150 – 159.

Nwana, H. (1996). "Software Agents: an Overview." Knowledge Engineering Review 11(3): 1-40.

Nygaard, V. and L. Fahrmeyer (2001). Choosing The Perfect Colors for Your Web Site, University of Arkansas at Little Rock. http://www.ualr.edu/factech/color_web_design.htm

Online, F. (2003). Families Online Web Page. http://www.lahaonline.com/People/index.htm.

Online, L. (2003). Laha Online Web Page. http://www.familiesonline.co.uk/main/topics.htm.

Paganelli, L. and F. Paterno (2002). Intelligent analysis of user interactions with web applications. Intelligent User Interface, San Francisco, USA, ACM Press: 111-118.

Pallant, J. (2001). SPSS Servival Manual. Buckingham, Open University Press.

Parush, A. (2001). "Usability Design and Testing." Interactions: 13-17.

Parzek, E. (1997). Shades of Success. Net Profit Magazine. http://www.trutlesweb.com/articles/color.html

Paterno, F. (2003). Tools for Remote Web Usability Evaluation. Human - Computer Interaction, Crete, Greece, Lawrence Erlbaum Associates.

Paymans, T. F., J. Lindenberg, et al. (2004). Short Papers: Usability trade-offs for adaptive user interfaces: ease of use and learnability. International Conference on Intelligent User Interfaces, Funchal, Madeira, Portugal, ACM Press: 301 – 303.

Pednault, E., N. Abe, et al. (2002). Sequential Cost-sensitive Decision Making with Reinforcement Learning. Conference on Knowledge Discovery in Data, SIGKDD, Edmonton, Alberta, Canada, ACM Press: 259- 268.

Perlman, G. (1996). Practical usability evaluation. Conference Companion on Human Factors in Computing Systems, Vancouver, British Columbia, Canada, ACM Press: 348 – 349.

Post, S. (2003). Saudi Post Web Page. http://www.saudipost.gov.sa/.

Preece, J., Y. Rogers, et al. (2002). Interaction Design: Beyond Human-Computer Interaction. United States of America, John Wiley & Sons, Inc.

Preece, J., Y. Rogers, et al. (1994). Human-Computer Interaction. Harlow, England, Addison-Wesley.

Ratner, J., E. Grose, et al. (1996). Characterization and Assessement of HTML Style Guides. Human Factors in Computing Systems, Vancover, Canada, ACM Press, 2: 115-116.

Rich, C. (1996). "Windows Sharing with Collaborative Interface Agents." ACM SIGCHI Bulletin 28(1): 70 - 78.

Rich, C., N. Lesh, et al. (2002). A Plug-in Architecture for Generating Collaborative Agent Responses. Autonomous Agent and Multiagent System, Bologn, Italy, ACM Press: 782-789.

Rich, E. and K. Knight (1991). Artificial Intelligence. New York, McGraw-Hill, Inc.

Rigden, C. (1999). "'The Eye of the Beholder'- Designing for Colour-Blind Users." British Telecommunications Engineering 1(17): 2- 6.

Rivlin, C., R. Lewis, et al. (1990). Guidelines for Screen Design. Oxford, Blackwell Scientific publications.

Rohn, J., J. Spool, et al. (2002). Usability in Practice: Alternatives to Formative Evaluations-Evolution and Revolution. Human Factors in Computing Systems, Minneapolis, Minnesota, ACM Press: 891-897.

Rojas, R. (1996). Neural Networks a Systematic Introduction. Berlin Germany, Springer-Verlag Berlin Heidelberg.

Rubin, J. (1994). Hand Book of Usability Testing How to Plan, Conduct Effective Tests. New York, John Wiley & Sons, Inc.

Rus, D., R. Gray, et al., Eds. (1998). Readings in Agents. Transportable Information Agents. San Francisco, Morgan Kaufmann Publisher Inc.

Russo, P. and S. Boor (1993). How fluent is your interface?: designing for international users. Human Factors and Computing Systems, Amsterdam, The Netherlands, ACM Press: 342 – 347.

Sánchez, J. A., J. J. Leggett, et al. (1994). "HyperActive: extending an open hypermedia architecture to support agency." ACM Transactions on Computer-Human Interaction (TOCHI) 1(4): 357 - 382.

Sawyer, P., A. Flanders, et al. (1996). Making a difference—the impact of inspections. Conference on Human Factors and Computing Systems, Vancouver, British Columbia, ACM Press: 376 – 382.

Scholtz, J., S. Laskowski, et al. (1998). Developing Usability Tools and Techniques for Designing and Testing Web Sites. Human Factor and the Web, Basking Ridge, NJ.

Seo, Y.-W. and B.-T. Zhang (2000). Learning User's Preferences by Analyzing Web-Browsing. Autonomous Agents and Multiagents, Bercelona, Spain, ACM Press: 381- 387.

Seo, Y.-W. and B.-T. Zhang (2000). A Reinforcement Learning Agent for Personalized Information Filtering. Intelligent User Interfaces, New Orleans, LA, ACM Press: 248-251.

Serban, G. (2003). "A new Reinforcement Learning Algorithm." Studia Univ. Babes-Bolyai, informatica XLVIII(1): 3-14.

Shapiro, D., P. Langley, et al. (2001). Using Background Knowledge to Speed Reinforcement Learning in Physical Agents. Autonomous Agents and Multiagents, Montereal, Quebec, Canada, ACM Press: 254- 261.

Shearin, S. and H. Lieberman (2001). Intelligent profiling by example. Intelligent user interfaces, Santa Fe, New Mexico, United States, ACM Press.

Shin, S., R. Gopal, et al. (2004). "Global Software Piracy Revisited." ACM Communications 47(1): 103- 107.

Shneiderman, B. (1998). Designing the User Interface: Strategies for Effective Human-Computer Interaction. Reading, MA., Addison-Wesley.

Shneiderman, B. (2003). Promoting universal usability with multi-layer interface design. ACM Conference of Universal Usability, Vancouver, British Columbia, Canada, ACM Press: 1- 8.

Sklar, J. (2000). Principles of Web Design. Cambridge, MA, USA, Course Technology Thomson Learning.

Sun, H. (2001). Building a Culturally-Competent Corporate Web Site: An Exploratory Study of Cultural Markers In Multilingual Web Design. SIGDOC, Santa Fe, New Mexico, ACM Press: 95-102.

Sutton, R. and A. Barto (1998). Reinforcement Learning: an Introduction. Cambridge, MA, MIT Press.

Terpstra, V. and R. Sarathy (1994). International Marketing. San Diego, Harcourt Brace College Publishers Fort Worth.

Terveen, L. G. and L. T. Murray (1996). Helping users program their personal agents. Human Factors and Computing Systems, Vancouver, British Columbia, Canada, ACM Press: 355 – 361.

TSB, L. (2003). Lloyds TSB Web Page. http://www.lloydstsb.com/.

Tumer, K., A. Agogino, et al. (2002). Learning Sequence of Actions in Collectives of Autonomous Agents. Autonomous Agetns and Multiagents, Bologna, Italy, ACM Press: 378- 385.

W3Consortium (2003). Evaluation, Repair, and Transformation Tools for Web Content Accessibility, http://www.w3.org/WAI/ER/existingtools.html

W3Consortium (2004a). HyperText Markup Language (HTML), http://www.w3.org/MarkUp/

W3Consortium (2004b). Leading the Web to Its Full Potential, W3Consortium, http://www.w3.org/

Walker, M. (2000). "An Application of Reinforcement Learning to Dialouge Strategy Selection in a Spoken Dialouge System for Email." Journal of Artificial Intelligence Research 12: 387- 416.

Walton, M. and V. Vukovic (2003). "Cultures, Literacy, and the Web: Dimentions of Information "Scent"." Interactions: 64-71.

Walton, M., V. Vukovic, et al. (2002). Development Consortium: 'Visual literacy' as challenge to the internationalisation of interfaces: a study of South African student web users. Human Factors and Computing Systems, Minneapolis, Minnesota, ACM Press: 530 – 531.

Ware, C. and W. Cowan (1990). "The RGYB Color Geometry." ACM Transactions on Graphics 9(2): 226- 232.

Wayner, P. (1995). Agents Unleashed a Public Domain Look at Agent Thechnology. Boston, AP Professional.

Webby The International academy of Arts and Science the Webby award judging criteria. http://www.webbyawards.com/main/webby_awards/criteria.html

Weeks, C. (1997). How to Design an Aesthetically-Pleasing ADA-Compliant Website. SIGUCCS XXV, Monterey, California, ACM Press: 319- 326.

Wilding, C. (1998). Practical GUI screen design: making it usable. Human Factors and Computing Systems, Los Angeles, California, ACM Press: 125 – 126.

Wlofman, S., T. Lau, et al. (2001). Mixed Initiatve Interfaces for Learning Tasks: SMARTedit Talkes back. Interactive User Interface, Santa Fe, New Mexico, ACM Press: 167- 174.

Wooldridge, M. and N. Jennings (1995). "Intelligent Agents: Theory and Practice." The Knowledge Engineering Review 10: 1-38.

Yap, M. and N. W. Keong (1999). Are Life-Like Characteristics Useful for Autonomous Agents. Autonomous Agents and Multiagent, Seattle, WA., ACM Press: 336-337.

Yeo, A. (1998). Cultural Effects in Usability Assessment. Human Factors and Computing Systems, Los Angeles, California, ACM Press: 74-75.

Zadeh, L. (1994). "Fuzzy Logic, Neural Network, and Soft Computing." Communications of the ACM 37(3): 78.

Zheng, Y. and S. Kiyooka (1999). Genetic Algorithm Applications, University of Victoria, British Coloumbia, Canada.